Tagged for Deletion: A Typological Approach to VP Ellipsis in Tag Questions

Craig Sailor
cwsailor@ucla.edu

UCLA Master's thesis
14 October 2009

Note to the reader:

Apart from a few organizational and typographical changes made in 2011, the version of the thesis you are reading here is largely unchanged from the original (as dated above). As a result, please note that this version does not fully reflect my current thinking on the topic, which differs in a few noteworthy, but relatively minor, ways.

Table of Contents

Acknowledgements

I would like to acknowledge and thank the members of my committee – Anoop Mahajan, Tim Stowell, and Carson Schütze – for tolerating incomplete drafts that should have been received sooner, for tolerating incomplete drafts that should have been shorter, and for tolerating incomplete drafts that should have been complete.

I would also like to recognize Sara Rosen for opening my eyes to syntax during my time at the University of Kansas.

Finally, I would also like to acknowledge and thank those closest to me for their unwavering support and encouragement.

Abstract

Following the results of a diverse typological study, I argue that "dependent" tag questions (henceforth "DTQs", defined below) are not unique to English, but can in fact be found in several languages, and they consistently pattern with instances of VP ellipsis (VPE) in each of those languages. I take these novel data to indicate that all DTQs are derived by VPE. I support this claim first in English, by showing DTQs to exhibit hallmark properties of VPE with respect to auxiliary stranding and licensing. Then, I show that DTQs behave like VPE in Taiwanese, Danish, Brazilian Portuguese, Scottish Gaelic, Persian, and likely Samoan. Thus, this study expands the empirical domain of ellipsis to include DTQs, and it also expands the empirical domain of DTQs to include languages other than English.

I define DTQs such as tag questions bearing a form-dependency on the clause that hosts them, including polarity sensitivity, "duplicated" TAM/verbal material, and pronominal reference. This contrasts with the non-DTQ type, which I do not discuss here.

English DTQs involve ellipsis. First, I show that English DTQs involve VPE by showing that they behave the same with respect to auxiliary stranding and licensing heads (Lobeck 1995). *Auxiliary stranding*: DTQs behave like VPE with respect to the auxiliaries they can and cannot strand, suggesting the same size of unpronounced structure in both. *Licensing heads*: the ellipsis site in DTQs always appears under a filled T^0 (which moves to C^0 in yes/no questions) or Neg^0. These licensing conditions are identical to those for VPE (Lobeck 1995).

Prediction: DTQs entail the availability of VPE. If a VPE approach to DTQs is correct, then I predict the following two linguistic universals to hold. *DTQ Implication*: if a language L has DTQs, then L also has VPE independently. *DTQ Generalization*: DTQs in L behave like

VPE in L with respect to the type(s) of stranded verbal material (auxiliary / main verb) and the presence of licensing heads in each.

Results: DTQs across languages attest the predictions. Each of the six languages in this study conforms to these universals, providing crosslinguistic evidence that DTQs involve VPE. *DTQ Implication*: each language with DTQs exhibits VPE in the canonical environments (e.g. in coordination, across S boundaries, etc.), verified as VPE (and not object-drop) with diagnostics from Goldberg (2005). *DTQ Generalization*: DTQs and VPE strand the same material in all the tested languages. When main verbs can survive VPE (*V-stranding VPE*: Goldberg 2005), the V also survives in DTQs. In the languages where VPE deletes V but strands auxiliaries, the same is true of DTQs: auxiliaries survive, but V does not. Brazilian Portuguese attests the full paradigm: VPE can strand either main verbs *or* auxiliaries, as can DTQs. Thus, VPE and DTQs across languages show the same sensitivity to the V-raising parameter, which follows if DTQs involve VPE.

This study has wide implications. First, it shows DTQs to be a crosslinguistically robust phenomenon. Second, it establishes that DTQs are derived by VPE, which has been assumed, but never supported. Third, it makes testable predictions phrased in the form of two linguistic universals, which can inform future work on VPE and DTQs. In addition to providing novel DTQ data, it also offers a rare look at VPE constructions in Taiwanese and Samoan; consequently, Taiwanese is revealed to be an aux-stranding language (a crosslinguistically rare property), while Samoan appears to be a V-stranding VPE language (similar to Swahili, Irish, Hebrew, and others). Finally, Brazilian Portuguese and Scottish Gaelic are shown to exhibit all of the relevant VPE properties already established in the literature for their genetic relatives, European Portuguese (Santos 2009) and Irish Gaelic (McCloskey 1991).

1. Introduction

The main goal of this thesis is to provide evidence from several languages showing that dependent tag questions involve "VP ellipsis" (henceforth, "VPE").[1] The source of this evidence is a novel typological study revealing several distributional and behavioral similarities between VPE clauses and tag question clauses across languages. Consider, for example, the similarities in "stranded" verbal material in Taiwanese and Persian tag questions and VPE:

Taiwanese tag question and VPE: auxiliary-stranding

(1) a. A-Ying **u** thak cit-pun che, kam **b-o** *Tag question*
A-Ying perf read one-class book Q neg-perf
"A-Ying read the book, didn't he?"

b. A-Ying **u** thak cit-pun che, A-Ha **b-o** *VPE*
A-Ying perf read one-class book A-Ha neg-perf
"A-Ying read the book, but A-Ha didn't"

Persian tag question and VPE: V-stranding

(2) a. Naysan ketaab-o na-**khoond**, **khoond** *Tag question*
Naysan book-obj neg-read read
"Naysan didn't read the book, did he?"

b. Naysan ketaab-o ba deghat **khoond**, Nasim ham **khoond** *VPE*
Naysan book-obj with caution read Nasim also read
"Naysan read the book carefully, and Nasim did (read the book carefully) too"

I take these data to be a small part of a larger empirical generalization that all dependent tag questions (defined below) are derived by VPE. I support this claim first in English, by showing that tag questions exhibit hallmark properties of VPE with respect to auxiliary stranding and licensing. Then, I show that tag questions behave like VPE in Taiwanese, Danish, Brazilian Portuguese, Scottish Gaelic, Persian, and likely Samoan. Thus, this study expands the empirical domain of ellipsis to include tag questions (which has been previously assumed but never

[1] The term "VP ellipsis" has become potentially misleading following recent work by Johnson (2004) and others arguing that canonical applications of ellipsis in the English verbal domain target *v*P, not VP. As no part of the forthcoming discussion will hinge on this detail, I use "VPE" as a theoretically neutral term throughout.

supported), and it also expands the empirical domain of tag questions to include languages other than English.

This wealth of empirical support leads me to develop an analysis of tag questions that relies heavily on independent properties of ellipsis, such as its syntactic and semantic licensing conditions (Lobeck 1995; Merchant 2001). Prior work on tag questions made valuable descriptive insights (Klima 1964, Huddleston 1970, McCawley 1988) as well as theoretical and conceptual insights (Bublitz 1979, Oehrle 1987, Culicover 1992, den Dikken 1995), and, while I adopt some parts of these past accounts, I largely argue in favor of a novel treatment of tag questions that avoids construction-specific stipulations. Instead, I claim that tag questions are simply adjoined yes/no questions that undergo VPE by way of their close semantic and pragmatic relationship to the clauses that host them. Thus, the syntactic derivation and semantic interpretation of tag questions follow entirely from independent grammatical principles (negation, interrogation, givenness, and the like), doing away with the need for a discrete "tag question construction" in the grammar, entirely.

1.1. *Background and definition*

While tag questions enjoyed a rich tradition in early generative linguistics, they have received little attention over the last decade. Many scholars have assumed tag questions to be the products of ellipsis for almost 40 years, but few empirical facts have been offered to support this assumption, and no principled effort to establish tag questions as involving ellipsis exists in the literature. As a result, the eruption of insightful work on ellipsis following recent observations by Johnson (2001, 2004) and Merchant (2001, 2004) has not included any discussion of tag questions whatsoever. I aim to remedy this shortcoming by providing data from tag questions that demand consideration given the unique challenges they pose for any theory of ellipsis.

The term "tag question" is ambiguous, so I will briefly define my usage of the term. Pre-theoretically, the term "tag question" typically refers to any sentence-final utterance whose presence has the effect of transforming a non-question into a question. Examining the inventory of tag questions in English, a binary classification arises: the forms of certain tag questions change according to the sentences they are construed with, while the forms of other tag questions never change. The first group represents *dependent tag questions*, and the second group, *invariant tag questions*. This distinction should be clear from the English examples in (3) and (4), which include tags from Canadian and British dialects:

Dependent tag questions

(3) a. [Sharon]$_i$ **could** probably pull a muscle doing that, **couldn't** she$_i$?
b. [Jeremy's restraining order]$_i$ **hasn't** already expired, **has** it$_i$?
c. [The fact that the police just arrived]$_i$ indicates that we ought to run, **doesn't** it$_i$?
d. [Your son]$_i$ **isn't** typically allowed to relieve himself in the dining room, **is** he$_i$?

Invariant tag questions

(4) a. Sally can't come because she's busy cleaning her dungeon, right?
b. Ron will be here soon with the crackers and spreadable meat, yes?
c. It's cold today, eh?
d. Tom's the one who likes that Swedish death-metal shite, innit?

For the remainder of this thesis, I restrict my discussion entirely to the dependent type. For brevity's sake, I refer to them simply as "tag questions" throughout, recognizing that an invariant type exists but is irrelevant to the present discussion.[2] I also exclude discussion of tag question intonation here (e.g. falling (rhetorical) vs. rising (information-seeking)), as well as "same-polarity" tags[3] (POS-POS: *Bill left, did he?*; NEG-NEG: **Bill didn't leave, didn't he?*). Each of these topics deserves a thorough treatment, which concerns of length and scope constrain me from applying here.

[2] See Appendix 1 for a survey of languages without dependent tag questions, and a sketch of their invariant type(s).
[3] See Appendix 2 for a survey of same-polarity tags excised from the typological study in section 4.

2. Prior approaches to the syntax of tag questions

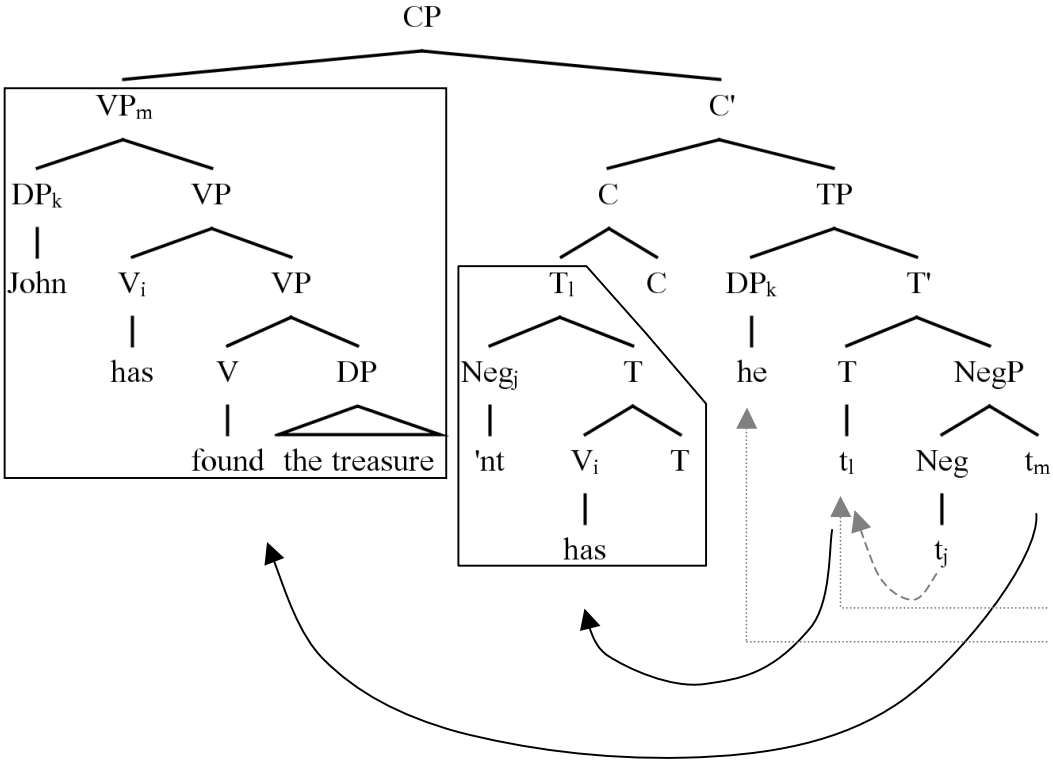
The form and structure of tag questions has been a recurring subject of debate since the development of Transformational Grammar. The nature of the relationship between the tag question clause and the clause that hosts it has been of particular interest, since tag questions superficially contain a portion of the antecedent clause, but not its entirety. Starting very early on, two conflicting theories emerged. The first held that tag clauses were derived by copying material/structure from a single clause (see Klima 1964 for the initial proposal; den Dikken 1995 offers a modern reinterpretation of the general idea); I will refer to this account as the *mono-clausal theory*. The alternative assumes that two clauses are involved – a host clause and a distinct tag question clause (see Huddleston 1970 for the initial proposal and Culicover 1992 for an updated implementation); I refer to this as the *bi-clausal theory*. While the two approaches differ on the syntactic status of the tag question itself (i.e. whether it is a clause), they are essentially alike in their derivation of the material within tag question: both approaches derive tags by some process of syntactic copying.

Early transformational analyses assume – often without elaboration – that tag questions are simply duplicates of the clauses that host them. Transformations apply to the duplicate that derive the appropriate surface form; however, the origin of the duplicate (i.e. how and why it appears in the first place) is rarely addressed. Indeed, almost all treatments of tag questions are primarily concerned with deriving the appropriate surface form of the tag clause, but very few address the means by which the tag clause appears to begin with. I describe the two most well-developed proposals in the subsections that follow, after which I argue that they are inadequate.

2.1. *Tags as pronounced traces of movement (Den Dikken 1995)*

Let us first consider the most restrictive implementation of copying: den Dikken's (1995) analysis (which is similar in spirit to Klima's 1964 original transformational analysis, although this is not cited in den Dikken's work). Seated firmly in Kayne's (1994) Antisymmetry framework, den Dikken's treatment of tags involves two fundamental steps: first, the tag material (subject, auxiliary, and optional negation) is copied into the T-layer; then, the VP moves to [Spec, CP] (which den Dikken equates with wh- movement). This fronted VP necessarily contains the in-situ subject, and auxiliary that had been copied into the T-layer; den Dikken implements this under a version of Spellout where the traces (copies) are pronounced out rather than deleted. The (simplified) structure is in (5) (adapted from his #20'):

(5) John has found the treasure, hasn't he?[4]



This proposal relies critically on the identity of the copied (i.e. moved) elements to their antecedents (i.e. spelled-out traces). The operations that drive movement (Move and, perhaps, Merge) blindly apply to eligible syntactic objects; alone, they are incapable of generating syntactic, morphological, or phonological differences between the moved object and its trace (with the exception of the pronominal subject of the tag, which den Dikken equates to a resumptive pronoun; see fn. 4 for discussion).

[4] According to den Dikken, the pronominal subject of the tag question is "a 'reduced' copy of the subject NP showing up in the internal subject position of the VP in [Spec, CP]. In essence, *he* is a resumptive pronoun; resumption can plausibly be looked upon as involving Chomskian copying" (den Dikken 1995: p. 10). He offers no further discussion of this analysis.

He also assumes that auxiliaries *and modals* are verbal heads that project VPs. Moreover, they are merged fully-inflected, so being severed from T^0 via VP-preposing does not affect their morphological form. Finally, den Dikken must claim that VP-internal subjects are merged into the specifier of the highest VP (which, in this case, is headed by an auxiliary). He offers little in the way of independent supporting evidence for these non-trivial claims.

Moving away from the specific mechanisms that den Dikken employs, we can consider the general spirit of the proposal, which is quite similar to Klima's 1964 transformational account. They argue that tag questions are built up from a single clause through a derivation that necessarily invokes movement (or reordering). Moreover, the material that composes the tag question – the subject and the auxiliaries – appears as the result of a copying operation presumed to exist independently the grammar. A strong prediction – one that seems trivially true, perhaps – arises from any copying approach: that is, the elements of a tag question should be duplicates of the elements in the host. We will see shortly, however, that this prediction fails to be confirmed in a non-trivial number of environments.

2.2. *Tags as complex anaphors (Culicover 1992)*

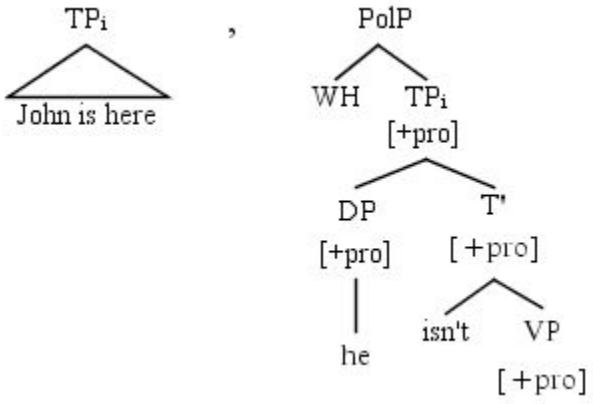
Before we see examples of such environments, though, we should consider a version of the copy-based approach that is somewhat less restrictive. While den Dikken (1995) assumes that tags are generated by blind syntactic copying (literally, the Move operation), others claim that tag clauses are built through some sort of anaphoric relation to the antecedent clause (implied in McCawley 1988; explicit in Culicover 1992). This allows for some flexibility in the material that appears in the tag, according to what can serve as an anaphor for the antecedent clause material. Culicover (1992) develops a theory that tag questions (and other constructions that resemble the remnants of CP- and VP-preposing he generally calls "Tags") are generated by a special "prosentential" clause that he calls pro-TP.[5] This pro-TP is coindexed with another TP in the discourse, and each node in the pro-TP receives its reference (and, in the case of the auxiliary verb(s) and agreement features, its morphophonological form) from this antecedent TP. Culicover implements this through an (optional) property of XPs he calls [+pro] (presumably

[5] Culicover (1992) uses IP rather than TP. Since no part of the theory hinges on this, I will use TP for concreteness.

short for "proform" and not the null nominal element *pro* of similar name). This property appears to be something like a phrasal feature (although this is not clear; see the next subsection for criticisms). This feature distributes through the entire structure of the pro-XP, such that each node dominated by XP obligatorily bears the [+pro] feature as well (see his #44 for a formal definition).

Each node bearing [+pro] must be bound in the traditional sense, similar to a regular pronominal. The tag's [Spec, TP] node, for example, bears [+pro] because it is dominated by a pro-TP; thus, the tag's [Spec, TP] is bound by the antecedent TP's corresponding [Spec, TP]. This binding operation results in the tag's [Spec, TP] being filled with a coindexed pronominal of the familiar type. The tag's pro-T^0 head is also bound by the antecedent T^0, resulting in what Culicover calls a "copy" (p. 207), and the pro-VP is valued as empty (i.e. elided). The structure corresponding to this analysis is represented in (adapted from Culicover's #45):[6]

(6) John is here, isn't he?



The subsequent subject-aux inversion seen in tag questions occurs, apparently, to satisfy the requirements of the WH "polarity" operator selecting the tag's pro-TP.

[6] The comma separating the two disconnected TPs in (6) is Culicover's (1992) notation. He does not offer an account of the structural relationship between the tag TP and the antecedent TP, which would seem to be necessary given his appeal to binding.

2.2.1. *A theoretical problem with the anaphoric approach*

If tags are complex anaphors (Culicover 1992), on the other hand, then a different set of problems arises. First, the introduction of a phrasal "[+pro]" feature requires a story to account for its apparently narrow distribution. If this feature is only licensed in tags, then it is ad hoc and makes no theoretical contribution (unless we assume that tags are a grammatical primitive, which Culicover specifically argues against). Likewise, the implementation of a [+pro] XP requires that anaphors exist for each corresponding node in the antecedent phrase. Assuming the notion of "anaphor" could somehow be expanded to include syntactic heads and projections (both lexical and functional), then any and all projections that could appear inside a [+pro] XP would require licit anaphors. This would entail a massive expansion of the lexicon to accommodate the requirements of a phrasal feature that otherwise does not appear to be widespread in the grammar.

2.3. *Summary*

The proposals in den Dikken (1995) and Culicover (1992) are perhaps the most thorough attempts at developing a syntax of tag questions since Klima (1964), and each has its advantages in capturing the data. At the same time, each proposal also has shortcomings that cannot be easily explained away. Moreover, both accounts rely on syntactic copying, despite that their general approaches differ substantially. In a mono-clausal approach, copying is necessary to keep the fronted constituent entirely intact while building the tag question material. The bi-clausal approach from Culicover (1992) also requires direct copying of at least T^0 (both its features and any auxiliary adjoined to it). In section 3.1, I explicitly argue against any sort of copying derivation for tags. First, though, I offer an alternative account in which tag questions are simply reduced question clauses that rely neither on copying nor a construction-specific anaphoric dependency.

3. Tag questions: pragmatics meets ellipsis[7]

If we do away with the assumption that tag questions are generated by some special operation (i.e. undeleted traces of head movement, anaphoric nodes, tag "copying," etc.), we can consider a theory that appeals only to independent principles of grammar – for example, adjunction, interrogation, polarity, and givenness. I propose that tag questions be thought of as regular yes/no questions that have undergone ellipsis of material that is e-given (Merchant 2001) from their antecedent clauses. Rather than being generated by some ad hoc operation of syntax, the content of the tag clause is dictated entirely by discourse as a necessary result of the pragmatics of tag questions.[8]

Informally put, in order to seek confirmation of a proposition (a primary function of tag questions, for example), that proposition must first be made explicit in the discourse. Requesting its confirmation can probably be done in a variety of ways, but one particular strategy available to English speakers is to question that proposition's logical opposite. As a consequence, this feeds ellipsis: the interpretation of that proposition is recoverable from the immediate discourse, which is the primary factor in the licensing of ellipsis (cf. Merchant 2001). A side effect of an ellipsis approach to tag questions is that tag clauses themselves are not treated differently in the syntax than any other clause that undergoes ellipsis (putting aside irrelevant structural differences; see 3.1). There is no grammatical mechanism that "generates" candidates for ellipsis or, more broadly, de-accenting; the clauses that undergo these operations do so because

[7] The data throughout this subsection are from American English. I note variations from these judgments in Canadian and British English where appropriate. The judgments reported here are my own, confirmed with several other native speakers of American English: thanks to Kyle Johnson, Robyn Orfitelli, Chad Vicenik, Byron Ahn, Jamie White, Nancy Ward, Peter Hallman, and anyone else I might be forgetting.

[8] Other scholars have recognized the need for a semantic/pragmatic approach to tag questions, including Huddleston (1970), Bublitz (1979), and Oehrle (1987). Culicover (1992) mentions (but does not define) a notion of "pragmatic consistency" to describe the relationship that holds between a tag clause and its antecedent. He argues that it could capture the patterns of polarity in tags without relying on syntax. I differ with Culicover in that he goes on to generate the remainder of the tag clause's material in the syntax (which I argue against in the previous section); however, I adopt his general appeal to pragmatics in the formulation of tag questions.

they meet the necessary conditions as a result of discourse. As a subtype of ellipsis clause, tag questions do not require any special treatment, either.

A consequence of this proposal is that there is no independent requirement on "tag question identity" that needs to be satisfied. Instead, the syntactic form of a tag question follows straightforwardly from pragmatics and ellipsis. An even broader consequence of this proposal is that a discrete tag question "construction" does not exist in the grammar, either. Instead, the set of data we call "tag questions" is simply the result of familiar grammatical operations acting in concert to generate reduced question clauses whose interpretations are dictated by discourse. Under this theory, the pragmatic functions traditionally associated to tag questions must thereby follow from the interacting semantics and pragmatics of these operations. Exactly how we arrive at this result – that is, why it should be the case that asserting a proposition and then questioning its polar opposite should yield a request for confirmation – I do not discuss here. Under the present proposal, however, this problem simply becomes a subset of the larger unresolved issues surrounding the interaction of these principles, such as how bias is derived from the combination of negation and interrogation in negative questions (see Romero & Han 2004 for identification and discussion of the open issues). Any progress toward answering these more general questions will therefore have direct implications for the theory of tag questions.

3.1. *Against copying: some exceptional tag question data*

As I mentioned at the end of section 2, prior analyses of tag questions (e.g. Klima 1964, den Dikken 1995, Culicover 1992) rely on a copying operation to capture some or all of the material appearing in tag clauses. This follows the intuition that tag clauses have an obligatory form-identity relation to their antecedents.

I argue against a copying approach to tag questions in this subsection by providing data showing that a requirement of strict form-identity does not exist: tag questions can and do differ from their antecedents phonologically, morphologically, and syntactically. These data cannot be straightforwardly explained by copying, suggesting that an alternative analysis – along the lines of the one I propose in the next section – is necessary.

3.1.1. *Coordinated-antecedent tag questions*

McCawley (1988: p. 482) observes that a coordination of clauses can collectively serve as the host for a single tag question clause:[9]

(7) John is drinking scotch and Mary is drinking vodka, aren't they?

Putting aside the details of coordinated-antecedent tag questions such as (7) for now (see 3.2), we can focus simply on the obvious problems they raise for a copy-based approach to tag questions. First, the tag's plural pronoun (*they*, above) cannot be a copy, since each of its overt DP referents (*John* and *Mary*) is singular. A movement account (à la den Dikken 1995) fails outright, since moving either of the antecedent subjects to generate their copies would constitute a Coordinate Structure Constraint violation. Even if both could somehow be copied into the tag's subject position and coordinated (which would also require explanation, since the subject coordination itself is not copied from anywhere in the derivation), the result would be the ungrammatical **...didn't he and she?* Likewise, the plural agreement on the auxiliary also cannot be copied (see 3.1.4 for a detailed discussion). Finally, the unpronounced material cannot be captured under an anaphoric approach, either, since the tag's antecedent is a coordination of

[9] McCawley judges (7) as "?". I find this sentence to be well-formed, as do the other speakers I consulted.

clauses, not VPs. More to the point, there is no conjunction of VPs anywhere in the derivation that could serve as the antecedent to a pro-VP of the sort Culicover (1992) proposes.

3.1.2. *The absence of variable binding and c-command*

Now, we turn to data that help reveal a structural fact of tag questions. Bound variable readings are known to require c-command, as in (8):

(8) a. [No sane person]$_i$ thinks that he$_i$ can beat me in a spitting contest.
b. [No man]$_i$ will admit that he$_i$ likes exfoliating microbeads in his bodywash.

Severing the c-command relationship yields ungrammaticality (or non-bound-variable readings):

(9) a. *[No sane person]$_i$ would ever drive in LA, and he$_i$ definitely wouldn't drive in Rome.
b. *If [no man]$_i$ will date my daughter, then he$_i$ is wise.

This relationship can be exploited to diagnose the structural relationship between tag clauses and their antecedents. That is, if the subject of a tag question can be interpreted as a bound variable, then c-command holds. Such a reading is impossible in tag questions:[10]

No c-command between antecedent and tag subject
(10) a. *[No man]$_i$ will ever scale Mt. Everest naked, will he$_i$?
b. *[No sane person]$_i$ rides motorcycles, does he$_i$?

This diagnostic indicates that antecedent subjects do not c-command into tag clauses. Rather than telling us what the structure of tag questions *is*, it tells us what the structure is *not*: namely, tag questions cannot be subordinate to their antecedents. This is consistent with McCawley's

[10] Some quantificational subjects sound better than others in these cases. I use the "no NP" type here to try to force a bound-variable reading. Curiously, the presence of a coreferential possessive object in the antecedent clause yields a much-improved reading:
i) ?[No girl]$_i$ hates her$_i$ own father, does she$_i$?
I leave this complication aside, since the simplest cases seem to indicate the absence of c-command.

(1988) argument that tag questions are adjoined to their antecedents, following examples from coordinated-antecedent tag questions (see 3.1.1).

Apart from being an important structural property of tag questions, the lack of c-command between the antecedent and the tag clause would seem to have implications for an anaphoric treatment of tag questions, as in Culicover (1992). He argues that pro-TPs such as tag questions are must be c-commanded by their antecedents (p. 213),[11] although he does not make the details of this requirement explicit. If we assume he means that the tag clause must be c-commanded by the highest projection of its antecedent, then this could fall out from an adjunction analysis; however, this structural relationship seems tenuous. Making use of the C-layer (by topicalizing a constituent, for example) would require projecting structure that would sever c-command between the antecedent TP and the pro-TP tag clause. If we assume that the tag clause is actually a pro-CP to avoid this pitfall, then any material in the antecedent CP (e.g. the topicalized constituent, the features on C^0, etc.) would be incorrectly copied into the tag clause.[12] As it stands, this anaphoric approach requires some adjustments to avoid these structural complications. On the other hand, the unavailability of c-command has no effect on the account of tag questions I offer here since they bear no strict syntactic relationship to their antecedents.

3.1.3. *Non-identical modals*

McCawley (1988: p. 482) also gives a handful of tag question examples whose modals are not obviously copied from their antecedents. For example, when certain modals that are unable to

[11] In an apparent contradiction, the structure of tag questions that Culicover (1992) offers (his #45, repeated in this thesis as (6)) does not include a structural relationship of c-command between the antecedent and the pro-TP.
[12] Under Culicover's (1992) approach, the alternative would be that the branches of the antecedent CP would incorrectly serve as antecedents for the [+pro] nodes within the tag clause CP. This would result in the copying of the antecedent C^0 into the tag, by analogy with T^0.

host cliticized negation (*may*, *might*, *ought*, etc.) would be forced to do so in a tag question environment, some other (non-identical) modal with an equivalent meaning appears instead:[13]

(11) a. We may have to work late, {*mayn't / won't} we?
b. We ought to leave now, {*oughtn't / shouldn't} we?

These examples pose a challenge for a copy-based approach to tags: simply, the copied forms yield ungrammatical sentences.

A copying account could perhaps be salvaged if the non-identical modals that appear in (11) were assumed to be alternative pronunciations of the antecedent modals when they combine with negation, akin to suppletion. Under this approach, though, the semantic similarity (near-synonymy) that holds between the non-negated modal and its "suppletive" form (i.e. **oughtn't* → *shouldn't*) is essentially a coincidence, since suppletive forms of a root are simply stored lexical entries associated with that root. Given that more than one alternative is possible for certain modals (e.g. **mayn't* → *won't* / *mightn't*), and the choices are subject to dialectal and speaker variation, it seems that semantics is playing a substantial role in choosing the viable alternatives. While this does not do away with the suppletion approach, it is exactly what we would predict if tag questions rely more on semantic and pragmatic coherence than form-identity.

3.1.4. *Non-identical auxiliaries: the syntactic autonomy of tag question clauses*

Other examples from McCawley (1988) that extend this conclusion exploit the optionality of singular/plural number in American English collective nouns denoting companies, organizations,

[13] For American English speakers, *ought* is at least slightly degraded by way of sounding literary and/or stilted even its positive form. Other deontic modals such as *shall*, *may*, and *must* – whose distributions have likewise become constrained over time in American English – trigger similar circumlocutions or "patches" in tag questions.

musical bands, etc. As full DPs, these typically trigger singular agreement on their verbs,[14] as we see in (12)-(14):

(12) a. The IRS **wants** my soul as collateral.
b. ??The IRS **want** my soul as collateral.

(13) a. N.W.A. **is** doing a Christmas album this year.
b. ??N.W.A. **are** doing a Christmas album this year.

(14) a. Google now **has** a fully-functional Terminator prototype.
b. ??Google now **have** a fully-functional Terminator prototype.

The (b) examples above are heavily marked in most varieties of American English,[15] suggesting that these types of collective nouns are singular, not plural. Interestingly, though, the morphology of pronouns coindexed with these collective nouns often displays a different pattern:

(15) a. The IRS$_i$ **wants** my soul as collateral, but **they**$_i$ can't have it.
b. N.W.A.$_i$ **is** doing a Christmas album after **they**$_i$ finish touring with Celine Dion.
c. Google$_i$ **has** a fully-functional Terminator prototype in **their**$_i$ Santa Monica office.

Despite triggering *singular* verbal agreement in the first conjunct, the collective nouns in (15) can (but need not) antecede *plural* coreferential pronouns in the second conjunct.

Appealing to this quirk of English, McCawley (1988) gives an example showing another way that a tag clause can look rather different from its antecedent, given here in (16) (example (a) adapted from his p. 482 #7; (b)-(d) are my own):

(16) a. IBM$_i$ **doesn't** make that model anymore, {**do** / ***does**} **they**$_i$?
b. The IRS$_i$ **wants** my soul as collateral, {**don't** / ***doesn't**} **they**$_i$?
c. N.W.A.$_i$ **is** doing a Christmas album this year, {**aren't** / ***isn't**} **they**$_i$?
d. Google$_i$ **has** built a fully-functional Terminator prototype, {**haven't** / ***hasn't**} **they**$_i$?

[14] Unless the collective noun bears overt plural morphology on the DP, in which case plural agreement is found:
i) The Beatle**s** {**are** / ***is**} underrated.

[15] Plural agreement is preferred in dialects of British English, however. Cheers to Tom Johnson for judgments.

The auxiliaries in these tag questions bear plural morphology that agrees with their plural pronominal subjects, whereas the full DP subjects in the antecedent clauses trigger singular agreement on their auxiliaries. This strongly suggests that the locality of agreement within the tag clause trumps the apparent identity relation that holds between a tag clause and its antecedent. If tag questions were derived by copying material from the full clause, then examples like (16) should be impossible: the pronominal subject should bear exactly the same phi-features (and thus number features) as the antecedent subject.

Neither of the analyses we saw in the previous subsection – nor any analysis relying on featural, morphological, or phonological identity between the tag and its antecedent – are capable of producing the grammatical examples in (16), meaning they undergenerate. On top of that, these analyses *are* capable of producing the *ungrammatical* examples, meaning they simultaneously overgenerate as well. An anaphoric approach might sidestep the problem of having a plural pronoun with a singular antecedent as some independent property of pronominalization in English (that is, whatever property would account for the data in (15)), but it could not explain the non-identical morphology on the auxiliary. This is because both T^0 and "AGR" are directly copied from the antecedent clause (Culicover 1992: p. 210). Likewise, a straight copying analysis falls well short of capturing examples like (16). In den Dikken's (1995) approach, for example, auxiliaries are merged fully inflected before being copied (via head movement); thus, a morphological mismatch between the tag's auxiliary and the one in the antecedent is completely impossible.

The data in (16) are especially important because they suggest that the tag clause is autonomous in its syntactic operations. That is, the above example shows that Agree applies independently to the tag question and the antecedent clause. In the tag question, this separate

application of Agree checks features (which can and do differ from those of antecedent clause) borne by lexical items merged from a unique numeration, not copied from elsewhere in the syntax. This is the same set of operations that applies to any full clause during a derivation.

3.1.5. *Non-identical subjects: "there" expletives*

Certain antecedent clauses allow (or force) a tag whose pronominal subject is not obviously coreferential with the structural subject of the antecedent clause. The examples in (17) contain non-specific subjects of copular clauses ((a)-(b), McCawley 1988: p. 482), along with a similar-looking case of locative inversion ((c), noted in Bowers 1976: p. 237). In these environments, the subject of the tag clause can be either a coreferential pronoun (which I find degraded) or an expletive *there*:[16]

Expletive "there" as the subject of a tag question
(17) a. Nothing$_i$ was broken, was {%it$_i$ / there}?
b. Six books$_i$ are on the shelf, aren't {%they$_i$ / there}?
c. In the garden is a beautiful statue$_i$, isn't {*it$_i$ / there}?

I discuss tag questions of this sort in more detail in 3.2. For our present purposes – ruling out copy-based approaches to tag questions – these data are plainly problematic for the analyses we saw in section 2: they have no obvious way of handling the unexpected appearance of the expletive *there* as the subject of the tag clauses in (17). In den Dikken's (1995) analysis, the tag question's subject is the head of a movement chain, but expletives famously appear only when subject raising has *not* occurred. We could perhaps concoct an "order of operations" account: if VP-preposing occurred before the subject could raise out of VP, then the subject would

[16] The acceptability of *it* and *there* in these tag questions is subject to a fair bit of speaker variation. I predict a bound reading of *it* to be unavailable here, given that antecedent material cannot c-command into the tag; see 3.1.2.
On the other hand, I predict an expletive interpretation of *it* here to be possible (if degraded). This follows from my analysis of reduced clefts in tag questions; see 3.1.6.

presumably not be in the right structural configuration to undergo copying into the tag (although this is not clear from den Dikken's account); this state of affairs would perhaps trigger *there*-insertion in the tag, instead. However, this approach would also overgenerate badly: it predicts the appearance of an expletive in every tag, which is clearly not attested.

3.1.6. *Non-identical subjects: "it" expletives*

A second class of non-referential subjects occurs in tag questions, as well – the expletive *it* shows up as the apparent subject of the tags in (18)-(20) (small caps indicate focal stress):[17]

(18) [Context: two people being introduced for the second time at a party]
a. TOM, wasn't it?
b. TOM, was it?

(19) [Context: two old friends recollecting stories from long ago]
a. NICK drove us home that night, wasn't it?
b. NICK wasn't driving us around that night, was it?

(20) [Context: scholars arguing about the authorship of a story attributed to Charlotte Brontë]
a. EMILY couldn't have written that story, could it have been?
b. Charlotte AND Emily could have written that story, couldn't it have been?

These sentences are less common when the "standard" tag is readily available, but they are still possible for many speakers, and again do not follow if the tag's subject (or auxiliary, in the case of (19)a for example) must be coindexed with overt material in the antecedent clause.

Section 3.2 contains a more detailed discussion of these examples. For now, we can simply take their existence as an additional empirical argument against any sort of copy-based approach to the derivation of tag questions, since, at the very least, the cleft-like structure in the tag (see 3.2) does not originate in the host clause in any of these examples.

[17] There is considerable speaker variation regarding the acceptability of these sentences, which seems to be categorical. Of the eight speakers I consulted, exactly half accepted them without difficulty (as I do), and the other half found them to be heavily degraded. This pattern of acceptability is interesting in its own right, and warrants more rigorous, experimental investigation. I leave this for future research.

3.1.7. *Summary*

The (mostly novel) data I present throughout this subsection raise several interesting questions, many of which I cannot address here. One thing that is certain, though, is that these data pose a challenge for analyses relying on syntactic parallelism, including Merchant's (2007 in progress, 2008b, 2009) implementation of a syntactic identity requirement on ellipsis, as well as den Dikken's (1995) copy-driven approach to tag questions. An approach that appeals to discourse and interpretation seems far more suitable (but, as it stands, is less restrictive) given the diversity of data showing non-identity between tag clauses and their antecedents. In each case of non-identity, a strong interpretational similarity holds. This intuition is already encoded in many theories of ellipsis, meaning tag clauses themselves can be captured without special treatment. Doing away with the assumption that tag questions bear strict syntactic identity requirements with their antecedents has important consequences for other works that have appealed to tags as diagnostics of various phenomena. Such appeals are widespread in the literature, particularly as attempts to determine the possible pronominal forms a constituent can take according to what shows up in the corresponding tag question's subject position (e.g. Mikkelsen 2005). Any conclusions predicated on the flawed assumption that tag questions are generated by syntactic copying need re-examination.

3.2. *Tag questions are VPE clauses*

The notion that tag questions are reduced interrogative clauses is certainly not new: this dates back to at least Huddleston (1970), and it is assumed in some form in Bublitz (1979), McCawley (1988), Culicover (1992), and many others. In fact, it has been all but taken for granted since Klima's (1964) original transformational analysis fell out of favor in the late 60s. Despite the near-ubiquity of this assumption, rather little empirical evidence has been offered showing that

tag questions pattern like independent instances of ellipsis. I aim to address this shortcoming here, by showing that tag questions behave like VPE clauses.

Diagnosing ellipsis in tag questions is, unfortunately, no simple matter. By their nature, tag questions provide frustratingly little material to manipulate: tag questions are maximally given in the discourse by definition, so the amount of new or contrastive material that can appear in the tag clause is severely constrained. As a result, some of the classical properties of VPE, such as the availability of sloppy identity, are impossible to test for in tag question environments. Likewise, diagnostics requiring c-command are of no use since such a relationship does not seem to hold between objects in the antecedent clause and the tag clause (see 3.1.2). The only obvious way to check whether tag questions involve ellipsis is to simply construct near-minimal pairs of the two phenomena, and see how they compare. As I show in the following subsection, this comparison yields a clear pattern: tag questions and VPE clauses have identical behavior with respect to the auxiliaries they can "strand" (leave undeleted), as well as the inflected material c-commanding the ellipsis site (Lobeck's 1995 "licensing conditions"). I take this as strong evidence that ellipsis is responsible for the reduction of the tag question clause. Later, in section 4, I present the results of a crosslinguistic study revealing that tag questions look identical to VPE clauses in precisely the same respect. Thus, the conclusion we reach for English in this section extends easily to a much larger and more diverse set of data that we will see later.

3.2.1. *Auxiliary stranding*

Lobeck (1987: p. 68-94) notes several distinct patterns of auxiliary stranding exhibited by VPE. I show that these patterns consistently appear in tag questions, as well. The complete range of "ellipsis licensing heads" appearing adjacent to VP ellipsis sites – for our purposes, the modals and auxiliaries (see section 3.3) – can also be found adjacent (before T-to-C) to the hypothesized

ellipsis site in tag questions.[18] Moreover, the same prohibition against including these heads in the VPE ellipsis site also exists in tag questions.

First, consider the modals:

Modals[19]

(21) a. Mister Ed couldn't read, but Arnold Ziffel sure could [~~read~~]. *VPE*
b. *Mister Ed couldn't read, but Arnold Ziffel [~~could read~~]. *VPE*
c. Mister Ed couldn't read, could$_i$ he t$_i$ [~~read~~]? *Tag question*
d. *Mister Ed couldn't read, (did) he [~~could read~~]? *Tag question*

Here, tag questions and VPE pattern alike: if there is a modal in the antecedent, it cannot be elided in the second clause. We can tell that the trace of the modal (presumably in T^0, before undergoing movement to C^0) is not deleted in the tag question clause in (21)c because in regular declaratives, modals are higher than uncliticized negation, and in tag questions, this negation survives:

(22) a. Most dogs can smell fear, but Sparky could not [~~smell fear~~]. *VPE*
b. Most dogs can smell fear, can$_i$ they t$_i$ not [~~smell fear~~]? *Tag question*

Thus, tag questions and VPE clauses pattern the same with respect to the distribution of modals.

Next, consider the behavior of perfective *have* in VPE and tag question environments. Similar to modals, perfective *have* can be stranded in both VPE and tag questions. Likewise, it

[18] I have excluded discussion of those licensing heads that take (small) clausal complements, such as *to*, since tag questions are generally simple clauses. This can be understood as a side effect of the pragmatics of tag questions, which dictates that the tag clause be maximally redundant (e-given) with its antecedent. A complex tag question is ruled out simply because it would require repeating a great deal of e-given material that could otherwise be elided, equivalent to the oddness of *Bill left, didn't he leave?*

[19] For brevity, I have only included *can*, but the same pattern holds of the other modals (*should*, *could*, *will*, etc.) as well. Certain modals (*shall*, *must*, *may*, etc.) do introduce complications, however; see fn. 13 for discussion.

must be stranded – it cannot be deleted – if it appears in the antecedent clause (in American English):[20]

Perfective "have"

(23) a. The pizza guy should have called, and the governor should have [~~called~~] too. *VPE*
b. #The pizza guy should have called, and the governor should [~~have called~~] too. *VPE*
c. The pizza guy should have called by now, shouldn't he have [~~called~~]? *Tag question*
d. *The pizza guy should have called by now, shouldn't he [~~have called~~]? *Tag question*

Again, tag questions and clauses undergoing VPE behave similarly with respect to auxiliary stranding: (23) shows that perfective *have* can and must be left out of the missing material in both VPE and tag questions. Although *have* appears to elide in (23)b, the only resulting interpretation does not match the tense/aspect of the antecedent. This suggests that the auxiliary responsible for this material (*have*) is never present in the VPE clause to begin with in (23)b.

Unlike modals and perfective *have*, though, non-finite forms of *be* can optionally elide in VPE environments. This optionality also shows up in tag questions:[21]

Non-finite "be"

(24) a. Dr. McCracken should be drinking, but I shouldn't (be) [~~drinking~~]. *VPE*
b. Dr. McCracken should be drinking by now, shouldn't he (be) [~~drinking~~]? *Tag question*

Furthermore, both VPE and tag questions allow multiple auxiliaries to be stranded, in which case the optionality of non-finite *be* persists:

Multiple stranded auxiliaries

(25) a. Phil should have been fired, but I shouldn't have (been) [~~fired~~]. *VPE*
b. Phil should have been fired, shouldn't he have (been) [~~fired~~]? *Tag question*

[20] Speakers of "Colonial" English (including the British and Canadian varieties) find (23)d acceptable, and even prefer it to (23)c. I constrict my discussion to American English, here.
[21] Lobeck (1987) and others report speaker variation on allowing *been* to elide in VPE contexts that I did not encounter except in speakers of Colonial English, which I do not treat here.

Unlike the perfective *have*, eliding non-finite *be* does not force a mismatch in tense/aspect between the VPE clause and its antecedent, suggesting it is present in the derivation, but unpronounced at PF.

Finally, the progressive form of *be* must obligatorily elide in VPE environments. Likewise, it cannot surface in tag questions either:

Progressive "be"

(26) a. Our hot dog vendor is being arrested, but our gyro guy isn't [~~being arrested~~]. *VPE*
b. *Our hot dog vendor is being arrested, but our gyro guy isn't being [~~arrested~~]. *VPE*
c. Our hot dog vendor is being arrested, isn't he [~~being arrested~~]? *Tag question*
d. *Our hot dog vendor is being arrested, isn't he being [~~arrested~~]? *Tag question*

Similar to non-finite *be*, the inclusion of the aspectual auxiliary *being* in the ellipsis site does not yield a mismatch in the aspectual interpretations of the ellipsis clause and its antecedent.

3.2.2. *Ellipsis licensing heads*

Lobeck's (1987, 1995) initial proposal of ellipsis *licensing* grew out of her observations that VPE only occurs when the ellipsis site is properly head-governed; in other words, the ellipsis site must be locally c-commanded by a lexical T^0 (at some point in the derivation). The class of licensing heads generally consisted of the elements that could appear in T^0, including the full complement of stranded auxiliaries and modals listed in 3.2.1, above. If the current proposal is on-track, meaning tag clauses are a subclass of VPE clause, then we expect the deletion in tag questions to be subject to the same licensing conditions as VPE. That is, we should not find any tag questions whose ellipsis sites are not immediately c-commanded by an overt lexical head.

The most obvious place to look is in cases where there is no overt auxiliary or modal in the clause. In VPE clauses, such environments trigger *do*-support, which satisfies the licensing requirement on VPE. *Do*-support also occurs in tag questions without filled T^0 heads:

T^0 Licensing of VPE and tag questions
(27) a. Although others *(did) [~~leave~~], Mary didn't leave.
b. Mary didn't leave, *(did) she [~~leave~~]?

This does not necessarily tell us anything about the licensing condition on tag question deletion, though: because tag questions are yes/no interrogative clauses, they require movement of material in T^0 to C^0. A classic distributional property of *do*-support is that it shows up any time the adjacency of V and T^0 are severed (Bobaljik 1994), which includes T-to-C environments. Although we cannot make the strong claim that tag clauses trigger *do*-support to satisfy the VPE licensing requirements, we can at least say that tag clauses do not violate the VPE licensing conditions in any obvious way.

There is at least one other environment in which ellipsis licensing seems to be identical between VPE and tag clauses. In addition to the material that can occupy T^0, Lobeck (1995) also argues that Neg^0 is a viable ellipsis licensing head, as well. This is testable in tag clauses, too. In regular yes/no questions where negation does not cliticize, the Neg^0 stays in-situ, despite the auxiliary material in T^0 moving to C^0:

Neg^0 in yes/no questions
(28) a. George does not wash his hands at reasonable intervals.
b. $Does_i$ George t_i not wash his hands at reasonable intervals?

Introducing VPE and tag questions into the data reveals that Neg^0 stays in-situ in these cases as well (unsurprisingly), meaning it is the VPE licensing head for both examples:

Neg^0 in VPE and tag questions
(29) a. Joy washes her hands compulsively, but George does **not** [~~wash his hands compulsively~~].
b. George washes his hands at reasonable intervals, $does_i$ he t_i **not** [~~wash his hands …~~]?

Example (29) shows that share Neg^0 as an ellipsis licensing head. Thus, while tag questions are not the easiest environments to diagnose licensing heads, it is at least clear that tags do not present any exceptional cases. While this is not positive support for the VPE analysis of tags, the licensing facts reinforce the theory by not challenging it.

3.2.3. *Revisiting some exceptional tag question data*

In this subsection, I take a closer look at some data presented earlier (in section 3.1), showing that such examples provide further support for a VPE analysis of tag questions.

3.2.3.1. *Ellipsis in coordinated-antecedent tag questions*

In 3.1.1, I presented data from McCawley (1988) showing that tag questions could essentially take split clausal antecedents (example below repeated from (7)):

(30) John is drinking scotch and Mary is drinking vodka, aren't they?

Given that the plural pronoun refers collectively to the subjects of each conjunct in the antecedent, McCawley takes this as evidence that the tag clause is not subordinate to its antecedent. Instead, he assumes an adjunction structure, whereby the tag clause is right-adjoined to the projection containing the coordinated clauses. I adopt his adjunction analysis here, recognizing that it cannot be empirically distinguished from silent coordination or parataxis (assuming reasonable analyses for each, although I am not aware of any in particular). This configuration predicts that antecedent material should not be capable of c-commanding into the tag clause, which is consistent with the discussion in 3.1.2.

Tag questions with coordinated antecedents exhibit a constellation of intriguing properties that McCawley actually does not discuss, as the above example in (30) is his only

mention of the phenomenon. I cover a few of their properties here. For brevity's sake, I refer to these phenomena as *coordinated-antecedent tags* (CATs).

First, in addition to allowing mismatching objects in the conjuncts, CATs also easily allow non-identical predicates:[22]

Non-identical predicates in CATs

(31) a. John sang and Mary danced a jig, didn't they?
b. Later tonight, John will read a book and Mary will watch TV, won't they?
c. John can't pick us up and Mary can't take us home, can they?

At this point, it is worth considering what the proposed ellipsis site would look like in these examples. I represent two possibilities for (31)a in (32):

Possible ellipsis sites for a coordinated-antecedent tag

(32) a. …didn't they [~~sing and dance a jig (respectively)~~] *distributive*
b. #…didn't they [~~(both) sing and dance a jig~~] *collective*

The only possible interpretation of (31)a is (32)a, where each referent of the tag's plural pronoun *they* is assigned to a predicate in a one-to-one fashion corresponding to their order in the antecedent clause.[23] I refer to this as the *distributive* reading. It is not possible to interpret the each referent as the subject of a coordination of the predicates, as in (32)b, which I refer to as the *collective* reading.

[22] Note that even when the predicates are identical, not all cases are acceptable:

i) *John$_i$ loves Mary$_j$ and she$_j$ loves him$_i$ too, don't they?

Presumably this is illicit because recovering the ellipsis site would require interpreting a reciprocal in the object position of the elided VP, as in "…don't they [~~love each other~~]?" I leave this open to future investigation.

[23] This is reminiscent of scope parallelism in VPE, noted by Fox (1995). He observes that VPE clauses and their antecedents must bear the same scope interpretations, even when each clause on its own is ambiguous. In this case, however, the parallelism only holds in tag questions: see the VPE example in (33).

If tag questions are in fact derived by VPE, then I expect any pattern available to tag questions to also be available in VPE. In this case, I expect VP ellipsis sites to be capable of taking coordinated antecedents, and they can:[24]

Coordinated-antecedent VPE

(33) a. John sang and Mary danced a jig. Come to think of it, Bill and Sue did, too.
b. …Bill and Sue did [~~sing and dance a jig, respectively~~], too *distributive*
c. …Bill and Sue did [~~both sing and dance a jig~~], too *collective*

Unlike tag questions, VPE clauses with coordinated antecedents are ambiguous: they allow either the distributive or the collective reading (although the former is preferable by way of being parallel to the antecedent). At present, I have no account for this interpretational asymmetry between tag clauses and VPE clauses. Still, we can take these data as broadly consistent with the analysis I pursue here, whereby tag clauses involve VPE.

3.2.3.2. *Ellipsis in tag clauses with expletive "there" subjects*

We saw examples of tag clauses containing expletive *there* subjects earlier, in 3.1.5 (example below repeated from (17)):

Expletive "there" as the subject of a tag question

(34) a. Nothing$_i$ was broken, was {%it$_i$ / there}?
b. Six books$_i$ are on the shelf, aren't {%they$_i$ / there}?
c. In the garden is a beautiful statue$_i$, isn't {*it$_i$ / there}?

We previously took such data as evidence against a copying analysis for tag questions. At this point, we can consider such examples in a bit more detail, and say something about their derivation.

[24] These data raise a substantial problem for VPE analyses requiring strict syntactic identity between the ellipsis site and its antecedent, à la Merchant (2008b). Simply put, there is no possible derivation for the ellipsis site in these data that is syntactically identical to that of its antecedent. See Baker (2007) for a possible workaround.

The most natural hypothesis that follows from these data is that the host clauses derive from those in (35) (i.e., from existentials):

(35) a. **There** was nothing broken, was {*it / there}?
b. **There** are six books on the shelf, aren't {*they / there}?
c. **There** is a beautiful statue in the garden, isn't {*it / there}?

This, in turn, yields a prediction: we only expect tags of the type in (34) to arise when their antecedents can independently appear as *there*-expletive constructions. This is confirmed:[25]

(36) a. *This was broken, wasn't there?
b. *There was this broken.

(37) a. *Six books fell off the shelf, didn't there?
b. *There fell six books off the shelf.

(38) a. *In the garden is my favorite statue, isn't there?
b. *There is my favorite statue in the garden.

Specific subjects and non-copular, non-existential clauses do not license the appearance of *there*. These facts suggest that tag questions are sensitive to other levels of representation beyond the surface antecedent they are construed with. This trait is not unique to tag questions, however: VPE can also make use of *there* when it does not appear overtly in the antecedent:

Expletive "there" in VPE: coordination
(39) a. I'm certain that nothing was broken, and Bill doesn't think {*it / there} was, either.
b. The picture showed six books on the shelf, but I didn't think {*they / there} were.
c. I told Bill that a statue was in the garden, but he didn't think {*it / there} was.

Generally, these data tell us that non-specific indefinites strongly disallow coreferential pronouns in most environments. The only available strategy in these cases, then, is to use *there*, despite its

[25] Locative readings of *there* should be ignored here, as we are only concerned with the expletive interpretation.

unknown status in the derivation. These also appear in disbelief questions and answers to yes/no questions (both of which I assume involve VPE):

Expletive "there" in VPE: disbelief questions
(40) a. Speaker A (confidently): Six books are on the shelf.
b. Speaker B (uncertain): Really? {*They / There} are [~~six books on the shelf~~]?

Expletive "there" in VPE: yes/no answers
(41) Q: Was anything broken?
A: No, {*it / there} wasn't [~~anything broken~~].
A: Yes, {*it / there} was [~~something broken~~].

Examples (39)-(41) show that this phenomenon is not limited to tag questions, but shows up in other VPE contexts, as well, again showing a tight correlation between tag clauses and 'regular' VPE clauses. Going further, perhaps it is not unreasonable to tie the availability of *there* in (34) and (39)-(41) to the recovery ("identification") of ellipsis, which is independently known to interact with its antecedent in certain unexpected ways (such as repairing island violations in sluicing; cf. Merchant 2008a). Thus, these data are probably best treated by a general theory of ellipsis, which I do not offer here.

Instead, assuming a bi-clausal analysis of tag questions, at least two basic approaches to these data are possible. The first possibility is derivational – namely, the tag question looks like it gets its reference from an existential copular construction because it actually does. In other words, we could assume some stage in the antecedent's derivation at which the expletive *there* is present in subject position, at which point the tag clause is created. While this approach is possible, it requires a few stipulations that currently lack independent evidence. On the other hand, having done away with a strict reliance on syntactic isomorphism for the formulation of tag clauses, a second possible approach would not need to make any special claims about the derivational stages of the antecedents in (34). Instead, this curious pattern would fall out as an

independent property of VPE and existential clauses (albeit one that requires further investigation).

3.2.3.3. *Ellipsis in tag clauses with expletive "it" subjects*

In 3.1.6, we saw examples of tag clauses that contained expletive *it* subjects (e.g. (19), reprinted below as (42)), despite that their host clauses contained referential subjects. We took this as a compelling argument against a copy-based approach to tag questions. Now, we will scrutinize their structure more carefully, and see that they provide additional arguments for the presence of VPE in the derivation of tag questions.

(42) [Context: two old friends recollecting stories from long ago]
a. NICK drove us home that night, wasn't it?
b. NICK wasn't driving us around that night, was it?

To understand the form of the tags in these examples, we must consider the salient interpretation of their antecedents. The focal stress indicated in each example is critical: without it, the above tags (and the others we saw in 3.1.6) sound quite odd. The necessity of contrastive focus in these examples follows if the tag clauses themselves are reduced ("truncated") cleft constructions – that is, if (42)a actually has the structure in (43):[26]

(43) NICK drove us home that night, wasn't$_i$ it t_i [~~Nick who drove us home that night~~]?[27]

These data strongly suggest that tag questions contain more structure than is pronounced, consistent with current theories of ellipsis, as well as reinforcing a pragmatic approach to tag

[26] See Mikkelsen (2007) for a review of the literature on truncated clefts, which she argues are neither clefts nor "truncated" by ellipsis. While the data I present here seem to pose challenges for her analysis – at least for the truncated clefts that appear in tag questions – I will save this discussion for future work.
[27] I take the necessity of contrastive focus in these examples to suggest that the elided material shares this information structure. If the tag were anything but a reduced cleft, such as a "wasn't it the case [$_{CP}$ …]" question, then the obligatory occurrence of contrastive focus is mysterious in these examples.

questions. Specifically, a reduced cleft is available as a tag because it can convey the same information structure as the antecedent clause, as clefted constituents are obligatorily focused.

This manipulation of information structure allows for a flexibility in tag questions that has been previously argued to be impossible.[28] With heavy focal stress, even objects and adjuncts can be the "subject" of tag questions using the truncated cleft strategy:

(44) a. Mark wasn't arrested that MONDAY, was it [~~that MONDAY that Mark was arrested~~]?
b. Mr. Nelson usually smokes opium BEFORE class, isn't it [~~BEFORE class that …~~]?
c. Doug went home with that girl BAMBI, wasn't it [~~that girl BAMBI who Doug …~~]?
d. The tiger mauled FIVE tourists, wasn't it [~~FIVE tourists that the tiger mauled~~]?

If these tag clauses really do contain reduced clefts, then they ought to be subject to the same restrictions as full clefts. This is attested in at least one environment. Consider the ungrammatical example in (45):

(45) *Sisyphus is HAPPY, isn't it?

The unacceptability of this example is predicted straightforwardly if the tag really is a reduced cleft: predicate adjectives cannot be clefted:

(46) *It is HAPPY that Sisyphus is.

If the tag questions of the type in (42)-(44) are *not* reduced cleft constructions, then the pattern in (45) is completely mysterious.

[28] Heycock & Kroch (2002) claim that reduced clefts are impossible in tag questions, citing the ungrammaticality of **I bought* THAT *book, isn't it*? However, this sentence is independently illicit because a) it has the pragmatically odd property of having a first person subject, and b) the cleft's copula fails to agree in tense with the antecedent (although it is not immediately clear why this should make a difference). Compensating for these issues, the example becomes much better: *Bill bought* THAT *book, wasn't it?*

Like the *there*-expletive data in 3.2.3.2, these *it* tag subjects can also appear in other VPE environments.[29]

Expletive "it" in VPE: coordination
(47) a. At the time, Bill thought NICK was calling, but it wasn't. It was just the police.
b. The newspaper reported that the tiger mauled FIVE tourists, but it wasn't – it was SIX.

Expletive "it" in VPE: confirmations
(48) A: ... then later that night, after Bill drove us home——
B: Wait a second – I'm pretty sure NICK drove us home that night.
A: It was [~~NICK who drove us home that night~~], you're right.

Expletive "it" in VPE: corrections
(49) A: …then all of a sudden, BILL walked in——
B: No it wasn't [~~BILL who walked in~~]. It was NICK.

Again, I take this distribution to indicate a relationship between tag questions and VPE. The fact that these truncated clefts can appear in apparent VPE environments indicates that it is not a tag-specific phenomenon. Instead, like the *there* expletive, it requires a broader description. Syntactic identity is not respected whatsoever between the ellipsis site and its antecedent, and yet ellipsis clearly applies; this conflict shows that considerations of interpretation (in this case, information structure) seem to trump those of surface form-identity, now familiar from our arguments against a copy-based analysis of tags.

3.2.4. *Summary*

The preceding discussion has provided several arguments for VPE in tag questions, including a set of behavioral similarities between tags and 'regular' VPE clauses. These similarities cannot be coincidental: they follow if tag questions do, in fact, involve VPE in their derivation. In the next section, I present a crosslinguistic study that bolsters this conclusion.

[29] The distribution of truncated clefts in VPE environments is not as wide as expletive *there*. I take this to follow from the rather delicate focus relationship that must hold between the cleft and its potential antecedent.

4. The crosslinguistic distribution of tag questions: a typological study

The remainder of this paper is dedicated to establishing tag questions as a crosslinguistic ellipsis phenomenon, aided by the results of an original typological study. Each of the six languages in the study has a tag formation strategy that bears an unmistakable similarity to its VPE strategy, as in (50):

Samoan

(50) a. na **faitau** e Sina le tusi, na le'i **faitau**-a
past read erg Sina the book past neg read-ES
"Sina read the book, didn't she?" (lit. "…not read?")

b. na **faitau** e Sina le tusi ananafi, a'o Ioane na **faitau**-a ananei
past read erg Sina the book yesterday but John past read-ES today
"Sina read the book yesterday, but John did today" (lit. "…but John read today")

The ellipsis-based account of tag questions I developed in section 3 accommodates this striking pattern straightforwardly: tags look like ellipsis because they *are* ellipsis. In what follows, then, the ellipsis approach to tags receives diverse empirical support. In extending the proposal to languages besides English, two predictions arise. First, I predict that each language with tag questions should exhibit evidence of a non-tag question verbal ellipsis phenomenon (assume VPE) in the canonical environments (e.g. in coordination). Second, I expect the tag questions in each language to pattern with these ellipsis phenomena in obvious ways, including, but not limited to, the size of the unpronounced structure involved and the verbal elements stranded by the deletion. These predictions are stated in the following pair of proposed universals:

Tag Question Implication

(51) If a language L has tag questions, then L also has VPE independently.

Tag Question Generalization

(52) Tag questions in a language L behave like VPE in L with respect to the type(s) of stranded verbal material (auxiliary and/or main verb) in each.

The implication in (51) is one-directional: a language with VPE will not necessarily have tag questions of the relevant type (e.g. Hebrew, Swahili, and others; see Appendix 1 for discussion). My goal in what follows is to show how these universals are exhibited by Taiwanese, Danish, Samoan, Persian, Brazilian Portuguese, and Scottish Gaelic.

4.1. *Introduction*

Prior work on tag questions has, to my knowledge, focused exclusively on English. Generally, the question of whether tags exist in other languages is not addressed; when it is, some works claim that tag questions are unique to English (Culicover 1992).[30] Incidentally, ellipsis was also claimed to be an English-only phenomenon for decades until important work starting in the early 90s began to reveal its influence in other languages (see e.g. McCloskey 1991, Doron 1999, and especially Goldberg 2005). Thus, the following typological study on tag questions is a natural extension of these works on the crosslinguistic distribution of ellipsis. While the data in this section show that tag questions are present in at least several distinct languages, they also reveal that the precise form of a tag question depends crucially on other factors in the language. Specifically, I argue that the surface variation in tags across languages derives from syntactic variation – specifically, *parametric* variation – with respect to the presence or absence of V-raising.

[30] To my knowledge, no prior work has considered tags from a crosslinguistic perspective. In fact, I know of only two sources that even contain examples of non-English tag questions, and neither work offers much discussion of the examples (as neither is concerned with tags specifically): see McCloskey (1991) for a single example in Irish Gaelic, and Mikkelsen (2005) for a handful of examples in Danish (limited to copular expressions).

4.1.1. *Motivating the study: a background on typology*

A central goal of linguistic typology is to recognize and classify structural properties of the world's languages. This sort of work is crucial to generative linguistics because it provides an empirical foundation, giving linguists a finer picture of what is possible in natural language. Most typological studies are extremely large-scale, appealing to data from hundreds of languages in an effort to eliminate statistical noise and converge on accurate crosslinguistic generalizations. Unsurprisingly, the number of these studies is rather small as a result. Baker & McCloskey (2007) argue for the utility and feasibility of a more constrained approach. In their recent exposition on the interaction of typology and theoretical syntax, they call for an expansion of diverse but *small-scale* studies (n = 5-10) to bridge the gap between the preponderance of single- and two-language studies featured prominently in the major publications in generative linguistics, and the massive typological studies that are too time-consuming to become widespread in the immediate future, despite their undeniable value. This alternative, which the authors call the "Middle Way," has the potential to make valuable contributions to both generative linguistics and typology by bridging the gap between the two in a way that is accessible to both sides. The modest study that I present here goes this Middle Way.

4.1.2. *Identifying the relevant parameter: V-raising*

Baker & McCloskey (2007) note that typological work stands the best chance at reliability "in areas where the space of possibilities is constrained fairly tightly by universal grammar" (p. 298). One such way of constraining the space is to appeal to the UG theory of parameters. For our purposes, a parameter can be defined as an irreducible point of potential variation across languages. *V-raising* is a classic example, and one that represents the simplest case by being

binary: languages are either V-raising or V-in-situ.[31] In a V-raising language (e.g. Hebrew, French, etc.), the highest verbal element (an auxiliary where present, otherwise the main verb) overtly raises out of the verbal domain to T^0, the locus of tense morphology. In Minimalist terms, main verbs in V-raising languages bear uninterpretable inflectional features that are said to be "strong," meaning they must be checked in a local (head-adjoined) configuration. On the other hand, V-in-situ languages like English never allow main verbs to raise to T^0 in simple declaratives;[32] instead, only the highest auxiliary can do so. In such languages, the main verb's uninterpretable inflectional features are said to be weak, meaning raising is not motivated.

The V-raising parameter directly interacts with ellipsis, which is best seen in an example. McCloskey (1991) convincingly argues that Irish Gaelic is a V-raising language, and yet it also has VPE.[33] The result is that V actually survives VPE ellipsis, while anything remaining in the verbal domain elides (example from McCloskey 1991: #29):

Irish Gaelic

(53) a. ar **cheannaigh** said teach?
C.int buy.past they house
Q: "Did they buy a house?"

b. creidim gur **cheannaigh**
I-believe C.past buy.past
A: "I believe they did." (lit. "I believe that bought")

McCloskey argues that examples such as (53) show that VPE in Irish Gaelic deletes the verbal domain, which includes any internal arguments (and, in the case of Irish Gaelic, the external argument as well; see McCloskey 1991: p. 263 for reasons why the subject stays "low" in Irish

[31] Barring complications involving, e.g., V2 phenomena, which are irrelevant to the current discussion.
[32] With the exception of main verb *be* and, possibly, *have*.
[33] He refers to the phenomenon as "SCE" (small complement ellipsis), noting that the ellipsis site is slightly larger than VP. As before, we will not be concerned with the precise size of the ellipsis sites in each of the languages treated here, as they are sure to vary a great deal, and diagnosing each one would easily exceed the intended scope of this paper. Instead, I continue to use the term VPE as a non-committal label for ellipsis of the verbal domain.

Gaelic). This general pattern of main verbs surviving ellipsis is called "V-stranding VP ellipsis" by Goldberg (2005), who goes on to describe its widespread distribution in the world's languages.[34] She lays out in exceptional detail a theory and corresponding set of diagnostics showing that VPE manifests in two fundamental varieties:[35] the V-stranding type of Irish, Swahili, Hebrew, etc., and the aux-stranding type of English (although Goldberg is only concerned with the former type in her study). I adopt Goldberg's (2005) classification and terminology for the organization of this section.

The languages in this study are consistent with the universals I propose in (51) and (52). The *Tag Question Implication* is upheld because each language in the study shows evidence of a tag question, as well as VPE. Likewise the *Tag Question Generalization* holds since the tag questions themselves pattern directly with VPE with respect to the verbal material excluded from the ellipsis site. What becomes clear is that the tested languages compose a proper subset of the typology of ellipsis in general, in line with our expectations. This is partially represented with respect to V-raising in Table 1:

Table 1: Ellipsis/tag question typology[36]

	V-raising	**V-in-situ**
VPE	V-stranding VPE Samoan, Persian, BP, Gaelic, Hebrew, Swahili… $[_{TP}$ (subj) V_i ~~$[_{VP}$ t_i (obj)$]$~~$]$	Aux-stranding VPE English, Taiwanese, Danish, BP… $[_{TP}$ (subj) aux ~~$[_{VP}$ V (obj)$]$~~$]$
Tags	V-stranding tags Samoan, Persian, BP, Gaelic $[_{CP}[Q]$ $[_{TP}$ (subj) V_i ~~$[_{VP}$ t_i (obj)$]$~~$]]$	Aux-stranding tags English, Taiwanese, Danish, BP $[_{CP}[Q]$ $[_{TP}$ (subj) aux ~~$[_{VP}$ V (obj)$]$~~$]]$

[34] Goldberg (2005: p. 212-213) notes that not all V-raising languages have VPE, most notably German, Dutch, and most of the Romance languages. It is not clear what the sufficient conditions are for a language to have VPE, but the presence of a verbal element locally c-commanding the ellipsis site appears to be a necessary condition. See Santos (2009: p. 98-100) for discussion and references.

[35] Toosarvandani (2009), using data from Persian, has subsequently argued for a third manifestation of verbal ellipsis – one where little-*v* is stranded, but the remaining verbal constituent elides (thus his term, "little-*v* stranding VP ellipsis"). This is reminiscent of the British English *do*-ellipsis construction discussed in Baker (2007), Haddican (2007), and Aelbrecht (2009).

[36] The simplified structures here ignore head order, meaning they are not accurate for Persian (SOV), for example.

Before getting to the data, though, I must mention some of the limitations of this study. The data I present here are original elicitations, but the elicitations themselves are extremely narrow in scope for each language, being mostly variations on *X read a book*. Most of these languages have other acceptable (sometimes even preferable) tag formation strategies; however, I only report those that fall into the "dependent" type defined in section 1.[37] Similarly, I do not provide analyses of the syntax of these languages, as doing so would far exceed the scope of this paper. Such work would be quite valuable, though, since various language-specific factors can directly affect the material that undergoes ellipsis and, by hypothesis, tag questions.[38]

Concerns of brevity, scope, and speaker availability also prevent me from applying the full range of VPE diagnostics to the constructions that "look" like VPE. I rely on prior authors' argumentation where available; otherwise, I supply at least one of Goldberg's (2005) diagnostics for distinguishing VPE from, for example, null anaphora (more on this in section 4.3). I must leave the application of the rest of her diagnostics, as well as any in-depth investigation of phrase structure, to future research on these languages. The goal of this study is simply to be a starting point in correlating tag questions and VPE across languages, not a comprehensive reference.

4.2. *Aux-stranding languages*

Languages whose tense and aspectual morphology can appear on verbal auxiliaries are not common crosslinguistically (Goldberg 2005: p. 21). These elements are merged higher than the main verb in the extended verbal projection, and most are outside the VPE ellipsis site (see

[37] See Appendix 1 for discussion of languages that lack tag questions of the dependent type.
[38] For example, Irish Gaelic subjects apparently fail to raise outside the ellipsis site in VPE, while other V-stranding VPE languages do not exhibit this property. See McCloskey (1991) for discussion.

3.2.1).[39] As a result, in languages with both an auxiliary system and a VPE process, VPE deletes the main verb, but strands higher auxiliaries (with some exceptions; again, see 3.2.1). Even in the absence of auxiliaries, main verbs in V-in-situ languages cannot survive VPE because they have no mechanism allowing them to escape the ellipsis site (i.e. they lack a strong uninterpretable inflectional feature that would allow them to raise outside the verbal domain to T^0, which auxiliaries typically have). In these circumstances, a pleonastic verbal element (a "dummy verb") often appears to express tense.

Of the six languages I found that had tag questions (excluding English), two of them exhibit aux-stranding as their only option in VPE:[40] Taiwanese and Danish. I describe their tag questions and VPE operations in the following subsections. I devote most of the discussion to Taiwanese for a few reasons. First, there is very little generative work on Taiwanese syntax, and none whatsoever that I could find on VPE (let alone tag questions); thus, the data I present is novel. On the other hand, Danish has been well-studied for some time, and its VPE (and even some of its tag question) phenomena have been noted elsewhere. Second, unlike Danish, Taiwanese is completely unrelated to English, so the existence of "English-like" tag questions there is all the more interesting.

[39] I take no stand in whether these elements head distinct projections (ProgP, AspP, PerfP, etc.), or whether they are simply iterations of higher verbal projections (perhaps v^0). See Aelbrecht (2009) for some discussion of the structure of these auxiliaries with respect to VPE (and other ellipsis phenomena).

[40] In addition, Brazilian Portuguese can strand auxiliaries (despite being a V-raising language), but it also has V-stranding VPE. For simplicity, I classify it as just a V-stranding language here, but see 4.3.4 for some discussion.

4.2.1. *Taiwanese*[41]

4.2.1.1. *Background*

Taiwanese (Southern Min, Sino-Tibetan) is an SVO language that can (but need not) mark tense/aspect with preverbal auxiliaries. One such auxiliary is *u* (roughly translating as *have*), a common marker of past/perfect:

(54) A-Ying **u** thak cit-pun che
A-Ying perf read one-class book
"A-Ying read the book"

In the past/perfect, sentential negation in Taiwanese typically relies on the presence of a particle *bo* (pronounced [bə25]),[42] which is in complementary distribution with the auxiliary *u*:

(55) A-Ying **bo** thak cit-pun che
A-Ying *bo* read one-class book
"A-Ying didn't read the book"

Although *bo* is traditionally said to be a negative marker, it seems likely that *bo* is actually the combination of a negative prefix *b-* and the auxiliary *u*, surfacing with a slightly centralized vowel quality. This analysis finds some support in data from the future/imperfective:

(56) a. A-Ying **e** thak cit-pun che
A-Ying imperf read one-class book
"A-Ying will read the book"

b. A-Ying **b-e** thak cit-pun che
A-Ying neg-imperf read one-class book
"A-Ying will not read the book"

[41] Linguistic resources for Taiwanese are not widespread. As a result, the transliterations and glosses throughout this subsection should be taken tentatively, as should any theoretical claims. I am deliberately ambiguous as to whether Taiwanese has tense (or just aspect). 多謝 to Grace Kuo for the data.
[42] Taiwanese is a tonal language. My transliterations leave out tonal information, though, as it is not crucial here.

In (56), we see that the future/imperfective auxiliary *e* becomes *be* when the sentence is negated, lending support to the idea that negative particles in Taiwanese are morphologically complex (see Cheng, et al. 1997 for a similar analysis and discussion).

4.2.1.2. *Taiwanese tag questions and VPE*

Taiwanese has tag questions of the relevant type. They minimally contain the Q-particle *kam* and an auxiliary, although the subject can optionally be expressed. If the antecedent contains an auxiliary, then this auxiliary also appears in the tag (similar to English):[43]

Taiwanese tag questions

(57) a. A-Ying **u** thak cit-pun che, (i) kam **b-o**
A-Ying perf read one-class book s/he Q neg-perf
"A-Ying read the book, didn't he?"

b. A-Ying **b-o** thak cit-pun che, (i) kam **u**
A-Ying neg-perf read one-class book s/he Q perf
"A-Ying didn't read the book, did he?"

Given that tag questions are present in Taiwanese, we expect VPE to be possible, and we expect the two to pattern alike. These expectations are met:

Taiwanese VPE

(58) a. A-Ying **u** thak cit-pun che, A-Ha ma **u**
A-Ying perf read one-class book A-Ha also perf
"A-Ying read the book, and A-Ha did too"

b. A-Ying **u** cim i e mama, A-Ha ma **u**
A-Ying perf kiss s/he poss mother A-Ha also perf
"A-Ying$_i$ kissed his$_i$ mother, and A-Ha$_j$ did {kiss his$_i$ mother / kiss her$_j$ mother} too"

In (58), everything below the aspectual auxiliary *u* is elided. The availability of sloppy identity in the elided constituent in (58)b shows that Taiwanese VPE behaves like English VPE.

[43] I note in section 3 that certain environments condition the appearance of a non-identical auxiliary. I have not checked for such environments in Taiwanese (or any other language I consider in this study), so I leave this open.

Moreover, the same amount of material seems to be missing from both the tag clauses in (57) and the VPE clauses in (58): main verbs, for example, cannot survive in either environment. I take this as evidence that Taiwanese tag questions are derived by ellipsis, and that Taiwanese obeys both the *Tag Question Implication* in (51) and the *Tag Question Generalization* in (52). The data in (57) reveal that Taiwanese tag questions are quite similar to those in English: the tag clause contains the same auxiliary as the antecedent clause. Under our bi-morphemic analysis of Taiwanese negation, the only relevant difference between the two clauses is polarity (because the auxiliaries are morphologically identical).

The appearance of *u* in the VPE and tag question clauses in (57) and (58) is worth discussing briefly. First, its appearance in tag clauses cannot be explained by any sort of copying operation discussed in section 2, since it appears even when there is no corresponding *u* in the antecedent clause:

(59)	A-Ying	thak-liao	cit-pun	che-a,	(i)	kam	**u**
	A-Ying	read-compl	one-class	book-part	s/he	Q	perf
	"A-Ying finished reading the book, did he?"						

In (59), the *-liao* suffix (which appears to encode completive aspect) appears in the antecedent, but the *u* auxiliary does not. Still, *u* appears in the tag clause here. Moreover, the imperfect auxiliary *e* cannot appear in this environment:

(60)	*A-Ying	thak-liao	cit-pun	che-a,	(i)	kam	**e**
	A-Ying	read-compl	one-class	book-part	s/he	Q	imperf
	"A-Ying finished reading the book, won't he?"						

These facts suggest that *u* is compatible with *-liao* in some way that *e* is not. I take this to instantiate the sort of pragmatic consistency we observed in English tag questions in section 3,[44] where non-identical auxiliaries (for example) could appear in tag clauses as long as their interpretations were near enough to those of the auxiliaries in the antecedent clauses. Whether this is the case, or *u* here is derived by some process akin to English *do*-support, the fact that *u* surfaces here supports Lobeck's (1995) original claim about the licensing of VPE in English. That is, she noted that VPE ellipsis sites were always locally c-commanded by a filled T^0. This same condition on ellipsis, then, seems to hold of Taiwanese, as well. The appearance of *u* in tags is independent support for my claim that they involve VPE, regardless of the process involved in its appearance (which clearly requires further investigation).

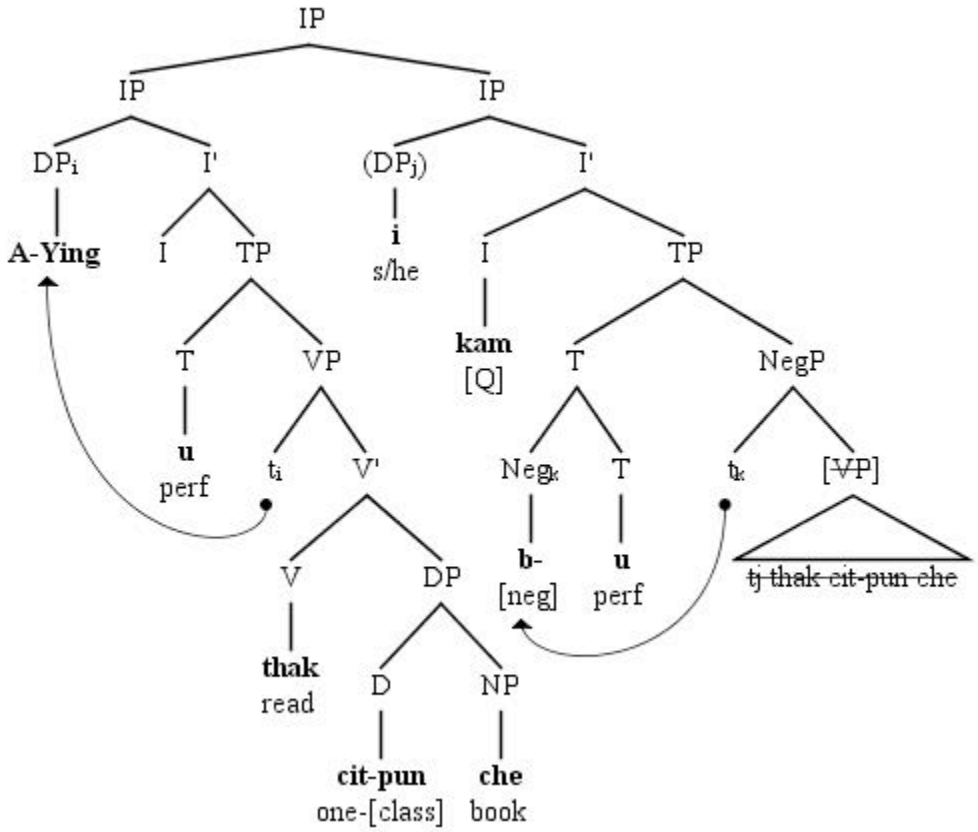
4.2.1.2.1. *A derivation for Taiwanese tag questions*

We are now in a position to consider a possible derivation for tag questions in Taiwanese. The status of the Taiwanese preverbal domain is not clear-cut, however. Huang (1991) takes for granted that subjects are in [Spec, TP] and assumes that *kam* (and Q-particles in other dialects of Chinese) are actually base-generated in T^0 (his "$INFL^0$"). This is somewhat unexpected, given the standard analysis from Chomsky (1995) and others that Q-particles typically occupy C^0. If Huang's analysis is on track, we must assume a "split-Infl" analysis, where *kam* is merged higher

[44] No part of the overall treatment of Taiwanese tag questions hinges on this, since the facts still support my claims.

in the structure than auxiliaries such as *u* and *e* (as Hagstrom 2006 notes).[45] With these assumptions in place, an illustration of (57), repeated here, is in (61):[46]

(61) A-Ying u thak cit-pun che, (i) **kam** **b-o**
A-Ying perf read one-class book s/he Q neg-perf
"A-Ying read the book, didn't he?"



[45] An alternative analysis is possible here, as Huang (1991) does not explain why *kam* must be in T^0. He argues that *kam* is the spellout of a Mandarin morpheme that drives (certain types of) A-not-A questions, which are formed by reduplication; however, he offers no arguments for its location in the T-layer. Presumably, its placement in T^0 would preclude reduplicating the subject, which is ungrammatical in A-not-A environments.

However, Huang's proposal for both Mandarin and Taiwanese suffers no ill effects if we assume that the morpheme is instead located in C^0 (consistent with other Q morphemes, à la Chomsky 1995). For Taiwanese, we only need to explain why subjects precede *kam* in the linear order, which follows if these subjects are actually topics (which are higher than C^0). Likewise, topics would not be reduplicated in Mandarin, so the A-not-A analysis stands.

[46] I represent the split-Infl here with IP and TP, but with no theoretical assumptions attached.

4.2.1.3. *Summary*

In this subsection, I present novel Taiwanese data that bear a noteworthy resemblance to English tag questions and VPE. I show that the analysis of English tags in section 3 can effortlessly extend to Taiwanese, whose tag questions are derived by VPE. This is a first step toward showing tag questions to be a crosslinguistic ellipsis phenomenon.

4.2.2. *Danish*[47]

4.2.2.1. *Background*

Danish (Indo-European: Germanic) is a V2 language with SVO word order, typically:

(62) a. Jasper læste bogen
Jasper read.past book.def
"Jasper read the book"

b. Jasper læste ikke bogen.
Jasper read.past not book.def
"Jasper didn't read the book"

Like Taiwanese and English, Danish is a language that makes widespread use of preverbal tense/aspectual auxiliaries. When present, these auxiliaries raise to T^0 to express tense, as expected (see Houser et al. 2009 for a summary of the Danish verbal system; example adapted from their #27):

(63) Mona og Jasper **havde** vask-et bilen
Mona and Jasper have.past wash-part car.def
"Mona and Jasper had washed the car"

[47] *Tak* to Troels Knudsen and, in particular, Line Mikkelsen for providing these data. I would not have thought to look for tag questions in Danish had I not seen her (2005) book on copular clauses, which contains a few examples that appeared to be of the relevant type.

These auxiliaries also undergo T-to-C movement in questions, similar to English (but unlike Taiwanese). Consider (64) (adapted from Houser et al. 2009 #31):

T-to-C in Danish questions (with auxiliary)

(64)	**havde**	Jasper	vask-et	bilen
	have.past	Jasper	wash-part	car.def

"Had Jasper washed the car?"

This same pattern arises in Danish tag questions, as we will see now.

4.2.2.2. *Danish tag questions and VPE*

When the antecedent of a Danish tag question contains an auxiliary, it also appears in the tag:[48]

Danish tag question (with auxiliary)

(65)	han	**havde**	slået	Udelvik,	**havde**	han	ikke
	he	have.past	slay-part	Udelvik	have.past	he	not

"He had slain Udelvik, hadn't he?"

Given the existence of tag questions in Danish, we expect to see evidence of VPE. Houser et al. (2008, 2009) list several examples, as well as supporting diagnostics (example from 2009: #27):

Danish VPE (with auxiliary)

(66)	Mona	og	Jasper	**havde**	vask-et	bilen,	eller	rettere	Mona	**havde**
	Mona	and	Jasper	have.past	wash-part	car.def	or	rather	Mona	have.past

"Mona and Jasper had washed the car, or rather Mona had"

Consistent with our predictions, VPE here mirrors the tag question exactly with respect to the stranded auxiliary (sans T-to-C). Thus, Danish tag questions are derived by ellipsis.

The Danish picture is not quite this simple, however. All of the examples we have considered so far have involved clauses with inflected auxiliaries. Mentioned previously, these auxiliaries move to T^0 to express tense, while the main verb stays in-situ, just like in English.

[48] This is a naturally-occurring example I found on a Danish blog, http://mandagsrollespil.blogspot.com/.

When no auxiliaries are present, however, Danish begins to look very different from English in certain environments, such as yes/no questions (Houser et al. 2009: #5b):

T-to-C in Danish questions (without auxiliary)

(67) **vaskede** Jasper bilen
wash.past Jasper car.def
"Did Jasper wash the car?"

Example (67) suggests that main verbs undergo T-to-C movement in questions in the absence of an auxiliary (cf. (64)), implying that they move to T^0. Danish, then, appears to be a V-raising language not unlike French: the highest verbal element – whether it is an auxiliary or a main verb – raises to T^0. French does not have VPE, however (Lobeck 1995).

Danish, on the other hand, clearly has VPE, as illustrated in (66) above (and see Houser et al. 2008, 2009). The languages that are now well-known as having both V-raising and a VPE operation – Hebrew, Irish Gaelic, Swahili, and so forth – all exhibit the same general pattern, laid out in great detail by Goldberg (2005): they exhibit so-called V-stranding VPE, because the main verb has raised outside the ellipsis site as a side-effect of its need to check inflectional features on T^0. We should expect Danish to fall into this class of V-stranding VPE languages, given that it has the necessary ingredients. Surprisingly, V-stranding is impossible in Danish:

Ungrammatical V-stranding in Danish

(68) a. *Mona og Jasper **vask-et** bilen, eller rettere Mona **vask-et** *VPE*
Mona and Jasper wash-part car.def or rather Mona wash-part
"Mona and Jasper washed the car, or rather Mona did"

b. *Jasper **læste** ikke bogen, **læste** han? *Tag question*
Jasper read.past not book.def read.past him
"Jasper didn't read the book, did he?"

Rather than stranding, the main verb actually elides in VPE and tag question clauses lacking an auxiliary, and a pleonastic verb *gøre* "to do" appears to express tense (see Houser et al. 2009 for extensive discussion of "*gøre*-support"):

Danish tag questions (without auxiliary: gøre-support)[49]

(69) a. Jasper læste bogen, **gjorde** han ikke
Jasper read.past book.def did.past he not
"Jasper read the book, didn't he?"

b. Jasper læste ikke bogen, **gjorde** han
Jasper read.past not book.def did.past he
"Jasper didn't read the book, did he?"

Now consider the continued resemblance of tags to VPE (this time, with respect to *gøre*):

Danish VPE (without auxiliary: gøre-support)

(70) Jasper læste bogen. **Gjorde** Mona også *VPE*
Jasper read.past book.def did.past Mona also
"Jasper read the book. Did Mona?"

Here is a brief review of the relevant facts and the questions they raise:

The Danish facts so far

(71) i. Danish is V-raising: aux moves to C^0 in questions: (64), otherwise V does: (67)
ii. Danish has tag questions and VPE, both of which:
a. Strand aux when one is present: (65)-(66)
b. Never strand V: (68)
c. Trigger *gøre*-support in the absence of an aux: (69)

[49] Although I have not included the data here, my investigation revealed Danish tag questions to be a licit environment for VP Pronominalization (VPP) (Houser et al. 2008). VPP in Danish is possible in nearly every canonical VPE environment (but not antecedent contained deletion, for example). In fact, Houser et al. (2009: fn. 9) note that VPE is actually dramatically less common than VPP in Danish. There are only a handful of environments where VPE is preferred (including questions, for example).

Interestingly, VPE is tightly constrained to these same contexts in Scottish Gaelic, as well. In many of the canonical VPE environments, Scottish Gaelic speakers prefer to use a pronominal rather than eliding the VP (Scottish Gaelic is a V-stranding VPE language, so V would survive). This pronominal could conceivably be a case of VPP (rather than the simple object pronominal it has been assumed to be). If this turns out to be the case, then the (near-)identical distribution of VPP and regular VPE in these two languages would be very interesting. Obviously, this requires more investigation. See section 4.3.2 for discussion of Scottish Gaelic.

The Danish puzzles

(72) i. If Danish is V-raising,
  a. Why does it lack V-stranding in tags and VPE?
  b. Why is *gøre*-support required in these environments?

I will briefly comment briefly on (72)i here. If we maintain Houser et al.'s (2009) V-raising analysis for Danish, then the impossibility of V-stranding VPE is completely mysterious: it implies that VPE environments alone are exceptional in disallowing V-to-T. However, if we assume that Danish main verbs do *not* undergo V-to-T movement *independent of movement to* C^0, then the facts become less challenging. In other words, the above contradiction is resolved if Danish has V-to-C in questions, but V-in-situ otherwise. The Danish V^0/v^0 can only move to T^0 on its way to C^0, and not otherwise. Assuming T^0 as the VPE licenser, the verb will still be in-situ at the point at which VPE applies (merger of T^0: Aelbrecht 2009),[50] correctly deriving the aux-stranding/*gøre*-support pattern: aux-stranding (with or without *gøre*-support) is the only possible ellipsis strategy simply because there is no other way to express tense. This analysis no doubt raises several problems, not the least of which is deriving the language's various V2 patterns, but the above facts leave few obvious alternatives. Maintaining this working hypothesis would require treating such environments as the exceptions, not the rules: in other words, the fact that the main verb occupies T^0 before moving to C^0 in questions (for example) is a side-effect of whatever drives movement to C^0 in questions in the first place (and the head movement constraint), *not* an independent property of Danish main clauses in general.

[50] If, however, VPE does not apply until T^0 is sent to Spell-Out (cf. Gengel 2007), and there are no other phase heads between C^0 and v^0, then V-to-C and VPE would be assessed within the same Spell-Out domain. Presumably, this is also the case for VPE and V-to-T movement, and yet V-to-T clearly applies, given V-stranding VPE. I leave these interesting issues for future research.

4.2.2.3. *Summary*

Danish clearly attests the implicational predictions from (51): because tag questions are present, VPE is also present; moreover, the two pattern identically with respect to the apparent size of the ellipsis site and the treatment of auxiliaries, particularly pleonastic *gøre*. In establishing these facts, I made a few additional observations along the way. In particular, I identified what appears to be a raising paradox: if Danish verbs always raise to T^0 in the absence of an auxiliary, then they should be capable of stranding when VPE has applied. The impossibility of such a configuration, as well as the obligatory appearance of a pleonastic verb in just such cases, indicates that V is staying within the verbal domain in declarative clauses. These facts cannot be reconciled without revisions to the current theory of Danish phrase structure, which I leave to future work.

4.2.3. *Summary of aux-stranding tag question languages*

Taiwanese and Danish both pattern quite similarly to English: each strands auxiliaries, but not main verbs, in VPE and tag questions. Likewise, when an auxiliary is not present in these environments, each language employs a pleonastic verb for the expression of tense. Assuming a V-in-situ analysis for all three languages allows for a uniform analysis of these phenomena. Thus, the first half of our typology is complete, and our predictions are attested for V-in-situ languages: the presence of tag questions indicates the presence of VPE, and the former patterns with the latter in every observable way. We turn our attention now to the V-raising languages, where the data look different, but the results are the same.

4.3. *V-stranding tag question languages*

Unlike V-in-situ languages that often rely on auxiliaries to handle the inflectional load, V-raising languages are much more common (although they too can allow auxiliaries, as in French and Brazilian Portuguese). The main verbs of these languages are the primary inflection-bearing elements; this is encoded formally by means of a strong uninterpretable inflectional feature on the verb that can only be valued by raising outside the verbal domain to T^0, the locus of inflection. This raising "saves" main verbs from VPE, whose ellipsis site is lower than T^0. The result is what Goldberg (2005) calls "V-stranding VPE".

In what follows, I show that Samoan, Brazilian Portuguese, Persian, and Scottish Gaelic have tag questions of the dependent type, and they all exhibit VPE. Most importantly, the tag question and VPE clauses in these languages exhibit the same V-stranding properties, which I take to be the final piece of evidence for the ellipsis analysis of tag questions I developed in section 3.[51]

V-stranding VPE languages pose a unique problem not necessarily present in the aux-stranding type, as pointed out in great detail by Goldberg (2005). That is, the only phonological effect of V-stranding VPE in the simplest cases is that the internal arguments are not pronounced. The same effect would be achieved in the case of object-drop. The problem is illustrated below: that is, the second conjunct of the sentence in (73) could have either possible structure in (74):

Brazilian Portuguese

(73) o Bruno não leu o livro [mas a Maria leu]
the Bruno neg read the book, but the Maria read
"Bruno didn't read the book, but Maria did"

[51] Several Tibeto-Burman languages (for which I do not currently have access to consultants) also appear to have dependent tag questions of the V-stranding type. For example, Singh's (1996) description of Meitei (Sino-Tibetan: India) strongly suggests that this language also features V-stranding tag questions. Further study of Meitei would no doubt be fruitful, especially to see what (if any) patterns of VPE arise in the language. I predict Meitei to have V-stranding VPE. Other candidates for V-stranding VPE languages from this sub-family include Tarao and Manipuri.

Possible structures: VPE vs. object-drop

(74) a. … o livro mas a Maria leu_k ~~$[_{VP}\ t_k$ o livro]~~ *VPE*

b. …[o $livro]_i$ mas a Maria leu_k $[_{VP}\ t_k$ $pro_i\ ]$ *object-drop*

The structures in (74)a and (b) are string-identical, meaning only structural diagnostics would be able to tease them apart. This is precisely what Goldberg (2005: p. 23-122) does in her treatment of V-stranding VPE: she establishes a collection of diagnostics that distinguish VPE structures from object-drop structures. Only when the object-drop structure in (74)b can be positively ruled out does she assume a VPE analysis for an ambiguous example such as (73).[52] Ruling out the possibility of object-drop typically requires exploiting language-specific restrictions on the phenomenon, as well as interpretational effects associated with VP-level adjuncts that go missing, etc. As I said in the introduction to this section, I cannot apply the full battery of these diagnostics here. Instead, I apply these diagnostics where I can, and leave systematic treatment of these data to future research.

4.3.1. *Scottish Gaelic*[53]

4.3.1.1. *Background*[54]

Scottish Gaelic (Indo-European: Celtic) is a VSO language:

(75) a. leugh Calum an leabhar sin
read Calum the book dist
"Calum read that book"

b. cha do leugh Calum an leabhar sin
neg past read Calum the book dist
"Calum didn't read that book"

[52] Because Goldberg (2005 p. 73-113) cannot rule out an object-drop analysis for putative VPE examples in Japanese and Korean, she argues that neither language has V-stranding VPE, contrary to prior claims. Incidentally, neither language has tag questions of the relevant type. See Appendix 1 for some discussion.

[53] *Mòran taing* to Muriel Fisher and Andrew Carnie for the data and discussion.

[54] In the examples I elicited, the past tense morpheme *do* only appears in negative and question clauses. I have nothing to say about this curious distribution.

Yes/no questions in Scottish Gaelic require a clause-initial Q-particle, *an*:

(76) an do leugh Calum an leabhar sin
Q past read Calum the book dist
"Did Calum read the book?"

Tag questions in the language follow the V-stranding pattern we have seen already for Samoan.

4.3.1.2. *Scottish Gaelic tag questions and VPE*

Tag questions in Scottish Gaelic minimally contain the Q-particle, *an*, as well as a main verb repeated from the antecedent clause:

Scottish Gaelic tag questions

(77) a. **leugh** Calum an leabhar sin, nach do **leugh**
read Calum the book dist neg.Q past read
"Calum read that book, didn't he?"

b. cha do **leugh** Calum an leabhar sin, an do **leugh**
neg past read Calum the book dist Q past read
"Calum didn't read that book, did he?"

The presence of an overt Q-particle in the tag clause (like Taiwanese) continues to demand an interrogative analysis of the tag clause. Following the *Tag Question Implication* in (51), the presence of tag questions should indicate the existence of VPE. Scottish Gaelic has VPE, to be sure, but the picture is not straightforward.

Andrew Carnie (p.c.) notes that Scottish Gaelic speakers generally reject VPE in many of the canonical environments associated with the phenomenon (such as coordination). In such environments, speakers prefer to use pronominalization instead:

(78) leugh Calum an leabhar sin agus leugh Mairi *(e) cuideachd
read Calum the book dist and read Mary it also
"Calum read the book, and Mary read it, too"

Carnie notes that VPE is possible in question environments, however.[55] This is also claimed for Irish Gaelic, a closely related language, by McCloskey (2005: #3):

Irish Gaelic VPE in yes/no questions

(79) a. **sciob** an cat an teireaball de-n luch
snatched the cat the tail from-the mouse
"The cat cut the tail off the mouse"

b. a-r **sciob**?
Q-past snatched
"Did it?"

For Irish Gaelic examples such as this one in (79)b, McCloskey argues that the question is formed by V-raising, followed by VPE. The presence of VPE in yes/no questions in Irish Gaelic supports Carnie's (p.c.) claim for Scottish Gaelic. Given that interrogative clauses are a licit VPE environment in Scottish Gaelic, it is not a coincidence that its tag questions behave like VPE: tag clauses in Scottish Gaelic are marked with *an* and *nach*. The presence of these Q-particles suggests that the tag clauses are interrogative, which in turn means that they are licit VPE environments. See McCloskey (1991) and Goldberg (2005) for extensive diagnostics verifying VPE in Irish Gaelic, including the VPE environments shared by Scottish Gaelic.

As further support for the existence of VPE in Scottish Gaelic, we see in (80) that minimal answers are possible:

Scottish Gaelic minimal answers

(80) Q: an do **leugh** Calum an leabhar sin
Q past read Calum the book dist
"Did Calum read the book?"

[55] At present, I have been unable to elicit an example. I appeal to Irish Gaelic to bolster Carnie's (p.c.) claims here.

A: **leugh**
read
"Yes, he read it"

A: cha do **leugh**
neg past read
"No, he didn't read it"

Minimal answers such as these in other languages are said to involve VPE, as in Persian (above), Portuguese (Santos 2009) and, particularly relevant here, Irish Gaelic (McCloskey 1991). Given all the similarities between Scottish and Irish Gaelic, it is not necessarily surprising to find that Irish Gaelic exhibits tag questions, as well (McCloskey 1991: #28):[56]

Irish Gaelic tag question
(81) **cheannaigh** said teach, nár **cheannaigh**
bought they house neg.Q bought
"They bought a house, didn't they?"

Rather than assuming that these similarities are coincidental, I take them to indicate the availability of VPE in Scottish Gaelic, consistent with the availability and behavior of tag questions in the language.

4.3.1.3 *Summary*

Here, I introduce Scottish Gaelic as another example of a V-raising language with tag questions. These tag questions are V-stranding, just like the (constrained) examples of VPE we find in the language, including the familiar-looking examples of minimal answers. As a side-effect of this investigation into Scottish Gaelic tag questions, Scottish Gaelic is revealed to have a great deal in common with Irish Gaelic: both languages have minimal answers, V-stranding VPE, and – most

[56] Seeing this isolated example in McCloskey (1991), which he mentions only in passing as part of a larger argument, was a big part of what gave me the idea to do this study in the first place; so, *go raibh maith agat* to McCloskey.

importantly to us – V-stranding tag questions. I now turn to Samoan, which exhibits all the same patterns we have seen in V-stranding languages so far, despite that it is entirely unrelated to them.

4.3.2 *Samoan*[57]

4.3.2.1 *Background*

Samoan (Austronesian: Polynesian) is an ergative VSO language:

(82) a. na faitau e Sina le tusi
past read erg Sina abs.the book
"Sina read the book"

b. na **le'i** faitau e Sina le tusi
past neg read erg Sina abs.the book
"Sina didn't read the book"

Yes/no questions are marked by falling intonation alone in Samoan, with no visible changes in word order or morphology. That is, the interrogative equivalents of (82)a and (b) would be string-identical, but uttered with falling pitch.

4.3.2.2. *Samoan tag questions and VPE*[58]

Samoan tag questions minimally contain a verb, which corresponds to the main verb in the antecedent clause:[59]

[57] *Fa'a fe tai* to John Fruean for the data, as well as Robyn Orfitelli and Hilda Koopman for discussion.
[58] Although I have not indicated it here, the tense marker *na* appears to be optional in tag questions. I have been unable to check whether it is also optional in VPE.
[59] Samoan tag questions exhibit the "uncertainty contour", an intonational pattern described by Orfitelli (2008) as appearing in utterances where the speaker is uncertain or doubtful about the truth of the proposition. Its appearance in tag clauses is not surprising, then: given that tag clauses deny the antecedent clause's proposition in order to seek confirmation of it, we fully expect the interpretation of the tag clause by itself to be "uncertain", while the antecedent is expected to be true. See section 3 for some discussion of the use and interpretation of tags.

Likewise, Mosel & Hovdhaugen (1992) note that the *-a* morpheme they gloss as "ES" ("ergativizing suffix") has several diverse functions, one of which is to mark "unexpected and inappropriate" action (p. 761). This particular function seems consistent with the "uncertainty contour" noted above, and with the pragmatics of tag questions in general. Its appearance in the VPE clause in (84) is due, perhaps, to this same property of the morpheme (given that the VPE clause is contrasted with its antecedent by way of *a'o* "but / whereas").

Samoan tag questions

(83) a. na **faitau** e Sina le tusi, na le'i **faitau**-a
past read erg Sina abs.the book past neg read-ES
"Sina read the book, didn't she?" (lit. "Sina read the book, not read?")

b. na le'i **faitau** e Sina le tusi, na **faitau**-a
past neg read erg Sina abs.the book past read-ES
"Sina didn't read the book, did she?" (lit. "Sina didn't read the book, read?")

This tag formation strategy is obviously different from the aux-stranding type exhibited by English, Taiwanese, and Danish (see section 4.2) since it involves repetition of the verb from the main clause rather than an auxiliary. This is consistent with V-raising, since the verb has moved outside the ellipsis site in the tag clause.[60]

Given that Samoan clearly has tag questions, we expect it to have VPE, following the *Tag Question Implication* in (51). This appears to be the case:

Samoan VPE

(84) na **faitau** e Sina le tusi ananafi a'o Ioane na **faitau**-a ananei
past read erg Sina the book yesterday but John past read-ES today
"Sina read the book yesterday, but John did today"

The data in (84) are similar to the tag questions in (83), since the main verb appears in both contexts. If tags really do involve VPE, then we expect such a similarity, assuming example (84) is a case of VPE.[61] Unambiguous confirmation of VPE in Samoan will strengthen the case for an

[60] Forthcoming work by Koopman (2009) argues that Samoan derives its V-initial order by VP-fronting, *not* V-raising. Specifically, Koopman argues for a predicate fronting operation that moves some remnant of the verbal complex containing the verb to a specifier position in the T-layer. The VPE/tag question data I present here may or may not have implications for this predicate-fronting analysis, which, as of now, has not been fully fleshed-out.

[61] At this time, I have been unable to secure a judgment for an example that would rule out a possible object-drop structure. Samoan independently prohibits the dropping of locative-marked internal arguments; thus, if an example such as (i) is possible in Samoan, it cannot be due to object-drop:

i) Na le'i va'ai Sina i le teine, a'o Ioane na va'ai
past neg see Sina loc the girl but John past see
"Sina didn't look at the girl, but John did"

If this example is judged grammatical, then it can only be by way of VPE. If the sentence is ungrammatical, on the other hand, then Samoan could potentially lack VPE, although more investigation would be necessary.

ellipsis analysis of tags, which already has diverse empirical support; however, this confirmation will have to wait for future work.

4.3.3. *Persian*[62]

4.3.3.1. *Background*

Persian (Indo-European: Iranian) is an SOV language:

(85) a. Naysan ketaab-o khoond
Naysan book-obj read
"Naysan read the book"

b. Naysan ketaab-o na-khoond
Naysan book-obj neg-read
"Naysan didn't read the book"

In colloquial Persian, yes/no questions bear declarative word order, but feature rising pitch.

4.3.3.2. *Persian tag questions and VPE*

Persian tag questions of the relevant type involve a repetition of the antecedent's main verb:

Persian tag questions
(86) a. Naysan ketaab-o **khoond**, na-**khoond**
Naysan book-obj read neg-read
"Naysan read the book, didn't he?"

b. Naysan ketaab-o na-**khoond**, **khoond**
Naysan book-obj neg-read read
"Naysan didn't read the book, did he?"

Persian tag questions appear to strand the main verb, consistent with the pattern we expect of a V-raising language. Given the *Tag Question Implication* stated in (51), we expect to see evidence of VPE. We do:

[62] *Merci* to Pariya Tehrani, Sahba Shayani, and Henry Tehrani for their judgments, and to Maziar Toosarvandani for extensive discussion. Persian has a rich written tradition, and its formal register varies greatly with its colloquial form. The data I present here are strictly from the latter.

Persian VPE

(87) Naysan ketaab-ro emrooz **khoond**, Nasim dirooz **khoond**
Naysan book-obj today read Nasim yesterday read
"Naysan read the book today, and Nasim did yesterday"

This pattern continues to resemble a V-raising language, since this looks like a prototypical case of V-stranding VPE. This is surprising, given properties of Persian I discuss shortly.

We can rule out the possibility that (87) is an instance of object-drop by appealing to a known V-stranding VPE diagnostic involving the interpretation of VP modifiers. Consider example (88)a and the interpretations in (b) and (c):

Ruling out an object-drop analysis for Persian VPE

(88) a. Naysan ketaab-ro ba deghat **khoond**, Nasim ham **khoond**
Naysan book-obj with caution read Nasim also read
"Naysan read the book carefully…

b. …and Nasim also did [read the book carefully]"

c. *…and Nasim also did [read the book]"

This example shows that the adverbial *ba deghat* is interpreted in the second conjunct, despite that it only appears in the first. Even if adverbials could drop independently in Persian, this would not explain why *ba deghat* is obligatorily interpreted in (88), since there is nothing requiring it to merge in the second conjunct in the first place. Instead, I argue along the lines of Doron (1999) and Goldberg (2005) for Hebrew, and Santos (2009) for European Portuguese, that this pattern can only be explained if VPE is applying in the second conjunct, and that recovering the elided verbal material under identity (or e-givenness, à la Merchant 2001) with the antecedent clause necessarily includes any adverbial modifiers. Thus, Persian can be independently established as having VPE, which fulfills the *Tag Question Implication* in (51).

The pattern we see in Persian VPE and tag question clauses resembles precisely the pattern we have seen for e.g. Scottish Gaelic and Samoan, both of which are uncontroversial V-raising languages. Crucially, however, Persian exhibits *none* of the telltale signs of V-raising: both adverbs and negation necessarily precede the V in unmarked orders. This is broadly consistent with SOV word-order typology.[63] Therefore, in the absence of positive evidence of a V-raising process, we must conclude that the Persian main verb does not occupy T^0 in declarative clauses. Given this result, we face an apparent paradox: the above examples suggest that Persian has V-stranding VPE (in both canonical and tag question contexts), yet it apparently lacks a necessary derivational component – V-raising – to derive the phenomenon.

A possible resolution to this paradox can be found in Toosarvandani (2009), which argues that Persian has an ellipsis operation he describes as *v*-stranding. That is, Persian is a language that makes robust use of light verbs (presumed to occupy v^0) to form complex predicates; and, as Toosarvandani convincingly shows, these light verbs can be left stranded when their VP complement undergoes ellipsis. The net effect is that these light verbs mimic English auxiliaries in being generated outside of the ellipsis site. With respect to ellipsis, then, the crucial difference between Persian and English reduces simply to the *size* of the ellipsis site: as we have seen, English elides nothing smaller than *v*P, whereas Persian apparently elides nothing larger than VP. Toosarvandani does not discuss simplex predicates in the language – that is, predicates lacking light verbs, having only main verbs; however, we can assume the same VP ellipsis operation is available in such cases by adopting the reasonable assumption that the Persian V raises to v^0 in the absence of overt material in v^0 (i.e. in simplex predicates). Thus, the examples we see above in (86)-(88) may indeed be called examples of V-stranding VPE, but

[63] Of which Persian is not a particularly well-behaved member: cf. the position of clausal complements (which follow the verb rather than preceding it), the position of complementizers (initial rather than final in the clause), etc.

only in the sense that the main verb in Persian has escaped a smaller ellipsis site (VP) by undergoing V-to-*v* movement. Empirical support for this set of circumstances in Persian tag questions comes from the fact that *v*-stranding VPE is also attested in tag questions, when their host clauses contain complex predicates (Toosarvandani, p.c.).

A non-trivial consequence of this reasoning – one that deserves some attention, but which I will have to leave aside here – is that Persian tag questions do *not* involve the same ellipsis operation seen in the other languages I discuss in this thesis (namely, *v*P ellipsis). They do, however, involve precisely the same ellipsis operation seen elsewhere *in Persian*. This bears on the formulation of the *Tag Question Implication* and *Generalization* stated earlier in (51)-(52): there, "VPE" must be interpreted as ranging over the language-specific predicate ellipsis operation in the language, regardless of the precise size of the ellipsis site. With this revision, tags in a language L are indeed still derived by "VPE" in L.

To conclude our discussion of Persian, we see that VPE is also instantiated in so-called *minimal answers* (Santos 2009), which are answers to yes/no questions consisting of a repetition of the main verb:

Persian minimal answers

(89) Q: Naysan ketaab-o **khoond** *(rising intonation)*
Naysan book-obj read
"Did Naysan read the book?"

A: chera, **khoond**
yes read
"Yes, he read it"

A: na, na-**khoond**
no neg-read
"No, he didn't read it"

While the availability of minimal answers in Irish Gaelic (McCloskey 1991), Hebrew (Doron 1999), and European Portuguese (Santos 2009) is well-established, they have not been previously noted for Persian, to my knowledge. As before, these can be distinguished from argument drop (in this case, subject- and object-drop), and represent further support for a VPE operation in Persian.

4.3.3.3. *Summary*

Persian obeys the *Tag Question Implication* and the *Tag Question Generalization* from (51) and (52): it has both tag questions and apparent VPE, and these two phenomena follow the same basic V-stranding pattern we expect of a language with V-raising. Persian is the only head-final language in this study, and yet the same patterns arise here that we will see in VSO languages (Scottish Gaelic and Samoan) and an SVO language (Brazilian Portuguese).

4.3.4. *Brazilian Portuguese*[64]

4.3.4.1. *Background*

Brazilian Portuguese (Indo-European: Romance; henceforth, "BP") is an SVO language. BP's syntax has a great deal in common with its European cousin, whose ellipsis properties have been well-noted in the literature; see Santos (2009: p. 21-112) for thorough discussion and many references. A basic picture of simple declaratives in BP with and without negation is in (90):

(90) a. o Bruno leu o livro
the Bruno read the book
"Bruno read the book"

b. o Bruno não leu o livro
the bruno neg read the book
"Bruno didn't read the book"

[64] *Valeu* to Tatiana Libman and Will Machado for the data and discussion.

Like Persian and Samoan, BP indicates questions with rising intonation on sentences bearing the same word order as simple declaratives.

4.3.4.2. *Brazilian Portuguese tag questions and VPE*

BP tag questions can be of the V-stranding type, in which case they minimally contain a verb found in the antecedent:

Brazilian Portuguese tag questions

(91) a. o Bruno **leu** o livro, não **leu**
the Bruno read the book neg read
"Bruno read the book, didn't he?"

b. o Bruno não **leu** o livro, **leu**
the bruno neg read the book read
"Bruno didn't read the book, did he?"

Given that tag questions are possible, we expect VPE to be, as well:

Brazilian Portuguese VPE (V-stranding)

(92) o Bruno **leu** o livro hoje, e a Maria **leu** ontem
the Bruno read the book today, and the Maria read yesterday
"Bruno read the book today, and Maria did yesterday"

Here, as before, we run into the possible object-drop confound noted in Goldberg (2005). Fortunately, though, VPE in European Portuguese (EP) has quite a robust literature verifying its existence, and given the extent of overlap in syntax of the two languages, the diagnostics established for EP can be carried over into BP. For example, Santos (2009) argues for VPE in EP using a similar VP modifier diagnostic as the one we applied to Persian in (88), above. This diagnostic also holds for BP, as well (example from Santos 2009: p. 27, adapted slightly for BP):

Ruling out an object-drop analysis for Brazilian Portuguese VPE (V-stranding)

(93) a. a Raquel não **deu** o livro para a mãe no Natal, mas a Ana **deu**
the Raquel neg gave the book to the mother on Christmas, but the Ana gave
"Raquel didn't give the book to her mother on Christmas…

b. …but Ana did [give the book to her mother on Christmas]"

c. *…but Ana did [give the book (to her mother)]"

The only possible interpretation for the second conjunct in (93) is one that recovers all of the material from the antecedent's verbal domain, including the internal arguments and the temporal adverb. Once again, a null anaphora account of these data cannot capture the obligatory interpretation of all these elements including *no Natal*. I take this to confirm the existence of VPE in BP, which further confirms the *Tag Question Implication* from (51). BP also shares the availability of minimal answers with EP:

Minimal answers in Brazilian Portuguese

(94) Q: o Bruno **leu** o livro *(rising intonation)*
the Bruno read the book
"Did Bruno read the book?"

A: **leu**
read
"Yes, Bruno read the book"

A: (não,) não **leu**
neg neg read
"No, Bruno didn't read the book"

Santos (2009) argues at length that the EP equivalents of these answers are derived by VPE, showing that they can be distinguished from cases of argument-drop, as they can in BP as well.

The VPE strategy we have seen so far for BP is clearly of the V-stranding variety. This is not a necessary condition on VPE in BP, however: it can also exhibit the aux-stranding variety (example (95)a from Santos 2009: p. 22 for EP, but confirmed for BP here):

Brazilian Portuguese VPE (aux-stranding)

(95) a. a Joana não **tinha** acabado o artigo mas a Teresa **tinha**
the Joana neg had finished the paper but the Teresa had
"Joana hadn't finished the paper, but Teresa had"

b. a Joana não **vai** ler o livro mas a Teresa **vai**
the Joana neg will read the book but the Teresa will
"Joana won't read the book, but Teresa will"

Fortunately for our theory, tag questions exhibit the same ability to strand auxiliaries:

Brazilian Portuguese tag question (aux-stranding)

(96) a. a Joana não **tinha** acabado o artigo, **tinha**?
the Joana neg had finished the paper had
"Joana hadn't finished the paper, had she?"

b. a Joana **tinha** acabado o artigo, não **tinha**?
the Joana neg had finished the paper had
"Joana had finished the paper, hadn't she?"

c. a Joana não **vai** ler o livro, **vai**
the Joana neg will read the book will
"Joana won't read the book, will she?"

d. a Joana **vai** ler o livro, não **vai**
the Joana will read the book neg will
"Joana will read the book, won't she?"

Thus, BP (and EP) can strand whatever is in T^0, since the highest verbal element is attracted there, regardless of whether it is an auxiliary or a main verb. The similarities between BP and EP, then, are not unlike those between Scottish Gaelic and Irish Gaelic. In fact, EP even has dependent tag questions of both the V-stranding and aux-stranding type, just like BP (naturally-occurring examples taken from Santos 2009: p. 146 fn. 6, and p. 161):

European Portuguese tag question (V-stranding and aux-stranding)

(97) a. **queres** jogar, não **queres**
want play neg want
"You want to play, don't you?"

b. nós também andamos zangadas, não andamos
we also are angry neg are
"We are also angry, aren't we?"[65]

The fact that both V-stranding and aux-stranding VPE are possible in both dialects of Portuguese reinforces the *Tag Question Generalization* in (52): whatever VPE is capable of in these languages, tag questions are capable of, as well.

4.3.4.3. *Summary*

Of all the languages included in this study, Brazilian Portuguese appears to be the most diverse in its VPE operations: it can strand whatever occupies T^0. This is just what one might expect of a V-raising language with auxiliaries and a VPE operation. Essentially, BP is what French and most of the other Romance languages would look like if they had VPE. The fact that tag questions also show generalized T^0-stranding in BP is the final piece of evidence that tag questions derive from ellipsis.

4.5. *Summary*

This small study makes two primary contributions. First, it establishes the *Tag Question Implication*, whereby the presence of tag questions in a language entails the presence of VPE. Second, it reveals that tag questions across languages adhere to the *Tag Question Generalization*, as they behave the same as the VPE operations that derives them. As I mentioned at the outset, this study raises a variety of questions, but its value is in confirming tag questions as deriving from VPE in several languages, and not just English.

[65] Santos (2009: p. 161) translates this example as, "We have been angry too, haven't we?"

4.5.1. *Directions for future research*

Although this study establishes the fundamental observations that hold between tag questions and VPE, there is still much that could be done. Foremost, an in-depth treatment of tag questions for any single language in this study (or otherwise) is sure to yield facts that escaped this survey. Empirical observations from tag questions must be taken seriously in the literature on ellipsis, since tag questions introduce a unique set of facts.

Any in-depth treatment would necessarily entail systematic application of Goldberg's (2005) diagnostics, which would secure a V-stranding VPE analysis for any cases that are still potentially unclear (such as Samoan). Likewise, the tag questions and VPE data I elicited in each language are severely constrained to the same basic sentences. A more thorough investigation would include a diversity of verbs (with respect to argument structure, for example), subjects (e.g. varying phi features: person, number, and gender), and auxiliaries (where possible). Examining the tag question clauses that arise as a result of such variations in the antecedent would no doubt be extremely fruitful.

It would also be worthwhile to thoroughly investigate whether the types of non-identical tags I introduce in section 3.1 can be found in other languages. For aux-stranding languages, are there any auxiliaries that lack a negative form? If so, what arises in the tag? In English, we see the appearance of non-identical auxiliaries whose meanings are nearly equivalent. We might imagine, though, that another language could employ the use of a pleonastic verb in such environments, instead. Given the *Tag Question Generalization*, I predict that such a pattern in either VPE or tag questions would necessarily implicate its appearance in the other. For V-stranding VPE languages, on the other hand, this raises an interesting question: are there

environments where VPE can strand non-identical Vs?[66] If so, we might expect the same pattern to appear in tag questions, assuming the pragmatic requirements of tags are respected. If VPE cannot strand mismatching Vs, then we do not expect such an environment to arise in tags. Either way, this is sure to tell us something about the role of syntactic identity in ellipsis.

Even though we have a much better understanding of tag questions at this point than we did to begin with, we still do not know what the sufficient conditions are for a language to have tag questions (although we know VPE is a necessary condition). We also do not know what the sufficient conditions are for VPE, either (cf. Goldberg 2005); however, recall that the *Tag Question Implication* is one-directional: if tags, then VPE. Given that we use the tag question as a starting point, we still need a theory that can explain why some languages have them and some languages do not. I leave this important question to future work.

I will briefly mention a direction that future research could take based on the contents of the Appendices. Although I have not treated invariant tags whatsoever in this thesis, I encountered several while I was gathering data. I grouped them crudely in Appendix 1 according to the status of VPE in the languages. Just as the preceding crosslinguistic study on dependent tags was entirely novel, so would such a study be for invariant tags, as well. While invariant tags superficially offer less to work with, given that their internal structure appears to be static, they seem to be far more widespread across languages, if not a true linguistic universal. I did not encounter a language that lacked a tag question of one sort or another. Moreover, there are clear patterns that emerge across the invariant tag types, as well: many include the same types of elements that arise in dependent tags, including question markers, negation, and some verbal-like element such as a copula (which, of course, does not vary according to its antecedent). As a closing remark on invariant tags, I suggest that any future work on the phenomenon should

[66] This is apparently possible in European Portuguese (Santos 2009: p. 57-58), but I have not investigated it for BP.

pursue whether an ellipsis account could be constructed for these, as well. Pursuing such an approach would likely entail positing ellipsis of a much larger size than VPE. Consider German's *nicht wahr* "not true", for example. It is conceivable, *prima facie*, that this (and other) invariant tags could involve ellipsis of an entire CP under identity with the CP it is adjoined to (e.g. [$_{CP}$ John left], [is it not (true) ~~[$_{CP}$ that John left]~~]?). If this were true, then the difference between dependent and invariant tags would almost be akin to the difference between VPE and, say, sluicing.

Likewise, Appendix 2 contains same-polarity tag questions, which I also did not treat in the body of this thesis. That being said, a universal pattern emerged in these data: no language allows both the tag and its antecedent to be negative simultaneously (with sentential negation). This would seem to derive from a semantic/conceptual prohibition. Specifically, I would guess that NEG-NEG is impossible because negative yes/no questions (as negative tags are, of course) presuppose the truth of a proposition, but this proposition is also negated in the antecedent clause. Thus, it seems to be a case of presupposition failure, and one that holds for all seven languages I considered. If this is on-track, then that indicates that the semantics of negative yes/no questions across languages share similar presuppositions. On the other hand, the availability of the POS-POS configuration is entirely mysterious. Approximately half the languages allowed them, and half did not. Given the unique set of pragmatic functions associated with POS-POS in English, I would not be surprised if this turned out to be a complicated problem.

5. Conclusion

This study has wide implications for the study of tag questions and VPE. First, it reveals that tag questions are a crosslinguistically robust phenomenon. Second, it establishes that tag questions are derived by VPE, which has been assumed, but never supported. Third, it makes testable predictions phrased in the form of two linguistic universals, which can inform future work on VPE and tag questions. In examining the crosslinguistic distribution of tag questions, it offers new insight into VPE constructions in Taiwanese and Samoan; consequently, Taiwanese is revealed to be an aux-stranding language (a crosslinguistically rare property), while Samoan appears to be a V-stranding VPE language (similar to Swahili, Irish, Hebrew, and others). Finally, Brazilian Portuguese and Scottish Gaelic are shown to exhibit all of the relevant VPE properties already established in the literature for their genetic relatives, European Portuguese (Santos 2009) and Irish Gaelic (McCloskey 1991).

It is my hope that this work will raise the empirical status of tags to a level where they are mentioned alongside other known VPE operations, such as minimal answers, alternative questions, etc. The tag clause's impossibly close ties with its antecedent introduce unique challenges for any theory of ellipsis, particularly given the small set of possible ways in which a tag can differ noticeably from its antecedent. These cases inspired the approach to tag questions that I offer in section 3, where I argue that tags are reduced yes/no questions that do not require the stipulative special treatment they had received in the past. I show that an account akin to that of VPE clauses gets more mileage out of the examples whose syntactic forms are clearly distinct from their antecedent, but whose interpretation is crucially similar.

References

Aelbrecht, L. 2009. *You have the right to remain silent: The syntactic licensing of ellipsis*. Doctoral Dissertation, Catholic University of Brussels, Brussels.

Baker, A. 2007. "Verb Phrase Ellipsis Resolution as a Side Effect of Discourse Coherence." University of Chicago ms.

Baker, M. and McCloskey, J. 2007. "On the relationship of typology to theoretical syntax." *Linguistic Typology* 11, 285-296.

Baltin, M. 2007. "Deletion versus pro-forms: A false dichotomy?" New York University, ms.

Bobaljik, J.D. 1994. "What Does Adjacency Do?" *MIT Working Papers in Linguistics* 22, eds. H. Harley and C. Phillips. 1-32.

Bowers, J. 1976. "On Surface Structure, Grammatical Relations, and the Structure Preserving Hypothesis." *Linguistic Analysis* 2, 225-242.

Bublitz, W. 1979. "Tag Questions, Transformational Grammar and Pragmatics." *Papers and Studies in Contrastive Linguistics* 9, 5-22.

Büring, D., and Gunlogson, C. 2000. "Aren't positive and negative polar questions the same?" UCSC/UCLA ms.

Cheng, L., Huang, J., and Tang, J. 1997. "Negative particle questions." *Microparametric Syntax and Dialect Variation, Current Issues in Linguistic Theory* 139, eds. James Black and Virginia Motapanyane. John Benjamins, Amsterdam, 41-78.

Chomsky, N. 1995. *The Minimalist Program*. MIT Press, Cambridge, MA.

Culicover, P. 1992. "English Tag Questions in Universal Grammar." *Lingua* 88, 193-226.

Dagnac, A. (to appear). "Modal Ellipsis in French, Spanish and Italian: evidence for a TP-deletion analysis." In *Selected Proceedings of LSRL 38* (38th Linguistic Symposium on Romance Languages, Urbana-Champaign, 4-6 april 2008), Benjamins.

den Dikken, M. 1995. "Extraposition as intraposition, and the syntax of English tag questions." Vrije Universiteit Amsterdam, ms.

Doron, E. 1999. "V-Movement and VP Ellipsis." *Fragments: Studies in ellipsis and gapping*, eds. Shalom Lappin and Elabbas Benmamoun. Oxford University Press, New York. 124-140.

Fox, D. 1995. "Economy and scope." *Natural Language Semantics* 3, 283-341.

Gengel, K. 2007. "Phases and Ellipsis." In *Proceedings of NELS 37*, eds. E. Efner and M. Walkow. Amherst: GLSA Publications.

Goldberg, L. 2005. *Verb-stranding VP ellipsis: A crosslinguistic study*. Doctoral Dissertation, McGill University, Montreal.

Haddican, B. 2007. "The Structural Deficiency of Verbal Pro-Forms." *Linguistic Inquiry* 38:3, 539-547.

Hagstrom, Paul. 2006. "A-not-a questions." *The Blackwell Companion to Syntax: Volume 1*, eds. Martin Everaert and Henk van Riemsdijk. Malden, MA: Blackwell. 174-213.

Heycock, C. and Kroch, A. 2002. "Topic, Focus, and Syntactic Representations." *Proceedings of WCCFL 21*, eds. L. Mikkelsen and C. Potts. Cascadilla Press. Somerville, MA. 141-165.

Houser, M., Mikkelsen, L., and Toosarvandani, M. 2008. "Verb Phrase Pronominalization in Danish: Deep or surface anaphora." *Proceedings of the Thirty-Fourth Western Conference on Linguistics* (WECOL34), eds. Erin Brainbridge and Brian Agbayani, 183-195.

Houser, M., Mikkelsen, L., and Toosarvandani, M. 2009. "Defective auxiliaries in Danish and English." University of California, Berkeley, ms.

Huang, J. 1991. “Modularity and Chinese A-not-A Questions.” *Interdisciplinary Approaches to Linguistics: Essays in Honor of Yuki Kuroda*, eds. C. Georgopoulos and R. Ishihara. 305-332, Kluwer Academic Publishers.
Huddleston, R. 1970. “Two Approaches to the Analysis of Tags.” *Journal of Linguistics* 6, 215-222.
Johnson, K. 2001. “What VP ellipsis can do, and what it can’t, but not why.” *The handbook of contemporary syntactic theory*, eds. Mark Baltin and Chris Collins. Oxford: Blackwell Publishers. 439-479.
Johnson, K. 2004. “How to be quiet.” *Proceedings of the Chicago Linguistic Society* 39.
Johnson, K. 2009. “Gapping is not (VP) ellipsis.” *Linguistic Inquiry* 40, 289-328.
Kayne, R. 1994. *The Antisymmetry of Syntax*. MIT Press: Cambridge, MA.
Klima, E. 1964. “Negation in English.” *The structure of language*, eds. Jerry Fodor and Jerrold Katz, 246-323.
Koopman 2009, Samoan ergativity as double passivization, UCLA ms.
Lobeck, A. 1987. *Syntactic constraints on ellipsis*. Doctoral dissertation, University of Washington.
Lobeck, A. 1995. *Ellipsis: Functional heads, licensing and identification*. New York: Oxford University Press.
McCawley, J. 1988. *The Syntactic Phenomena of English*. University of Chicago Press, Chicago.
McCloskey, J. 1991. “Clause Structure, Ellipsis and Proper Government in Irish.” *Lingua* 85, 259-302.
McCloskey, J. 2005. “Predicates and Heads in Irish Clausal Syntax.” *Verb First: On the Syntax of Verb-Initial Languages*, eds. Andrew Carnie, Heidi Harley, and Sheila Ann Dooley. John Benjamins, 155-174.
Merchant, J. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
Merchant, J. 2003. “Remarks on stripping.” University of Chicago ms.
Merchant, J. 2004. “Fragments and ellipsis.” *Linguistics and Philosophy* 27, 661-738.
Merchant, J. 2007 (in progress). “Voice and ellipsis.” University of Chicago ms.
Merchant, J. 2008a. “Variable island repair under ellipsis.” *Topics in ellipsis*, ed. Kyle Johnson, 132-153. Cambridge: Cambridge University Press.
Merchant, J. 2008b. “An asymmetry in voice mismatches in VP-ellipsis and pseudogapping.” *Linguistic Inquiry* 39, 169-179.
Merchant, J. 2009 (to appear). “Ellipsis.” *Syntax: An international handbook*, eds. Artemis Alexiadou and Tibor Kiss. Berlin: Mouton de Gruyter.
Mikkelsen, L. 2005. *Copular Clauses: Specification, Predication and Equation*. Linguistik Aktuell 85, Amsterdam: John Benjamins.
Mikkelsen, L. 2007. “On so-called truncated clefts.” *Kopulaverben under Kopulasätze: Intersprachliche und Intrasprachliche Aspekte*, eds. Ljudmila Geist and Björn Rothstein, Tübingen: Niemeyer Verlag. 47-68.
Mosel, U. and Hovdhaugen, E. 1992. *Samoan Reference Grammar.* Scandinavian University Press, Oslo.
Ngonyani, D. 1996. “VP Ellipsis in Ndendeule and Swahili Applicatives.” *Syntax at Sunset: UCLA Working Papers in Syntax and Semantics* 1, eds. Edward Garrett and Felicia Lee. 109-128.

Ó Siadhail, M. 1989. *Modern Irish, Grammatical Structure and Dialectal Variation*. Cambridge University Press, Cambridge, England.
Oehrle, R. 1987. "Multi-dimensional categorial grammars and linguistic analysis." *Categories, Polymorphism and Unification*, eds. E. Klein and J. van Bentham, Center for Cognitive Science, University of Edinburgh, Edinburgh, 231-260.
Orfitelli, R. 2008. "Intonational phonology of Samoan interrogatives." UCLA, ms.
Reese, B. 2007. "Bias in Questions." Doctoral dissertation, UT Austin.
Romero, M. and Han, C. 2004. "On negative yes/no questions." *Linguistics and Philosophy* 27, 609-658.
Ross, J.R. 1967. *Constraints on Variables in Syntax*. Doctoral dissertation, MIT.
Santos, A. L. 2009. *Minimal answers: ellipsis, syntax, and discourse in the acquisition of European Portuguese*. John Benjamins, Amsterdam.
Singh, C.Y. 1996. "Tag questions in Meiteilon." *Mon-Khmer Studies* 25, 29-38.
Toosarvandani, M. 2009. "Ellipsis in Farsi complex predicates." *Syntax* 12, 60-92.

Appendix 1: Languages without dependent tag questions

In searching for languages with tag questions, I was able to rule out several candidates after consulting with native speakers.

I include a very brief description of all non-dependent tag question strategies I encountered for each language. The list of strategies should not be taken to be exhaustive, since my primary goal was to establish the (non-)existence of dependent tags in the languages I examined, rather than compiling a complete list of the possible non-dependent strategies.

Likewise, the listed VPE statuses given below are based solely on whether I was aware of any published accounts of such an operation in these languages (thus, "unknown" means "unknown to me"). Clearly, these should be taken tentatively, as I have not conducted any kind of systematic investigation verifying these facts.

Table 1: Languages lacking dependent tag questions[67]

VPE languages	Non-VPE languages	VPE status unknown
Hebrew (y/n words; "right?"; "what?")	French (y/n words; "is it not?")	Hindi ("not?")
Swahili ("is not?")	Italian ("not true?")	Zapotec (question particle)
Vietnamese ("yes no?")	German ("not true?"; "or?")	Kaqchikel (question particle)
Malagasy ("or not that?"; "that?")	Dutch ("or not?"; "or yes?")	Swedish ("or how?")
	Korean ("so?")	Cantonese (A-not-A)
	Japanese (copula + particle)	

[67] Thanks to Roy Becker, Harold Torrence, Khanh Nguyen, Ed Keenan, Dominique Sportiche, Flavia Adani, Daniel Büring, Tim Stowell, Hilda Koopman, Jieun Kim, Tomoko Ishizuka, Anoop Mahajan, Felicia Lee, Ana Lopez de Mateo, Ingvar Lofstedt, and Lawrence Cheung for their judgments and/or help in my collection of these data.

Appendix 2: Patterns in same-polarity tags

Taiwanese: *NEG-NEG / POS-POS

Same-polarity tags in Taiwanese

(98) a. *A-Ying **bo** thak cit-pun che, (i) kam **bo**
A-Ying perf read one-class book s/he Q neg-perf
"A-Ying didn't read the book, didn't he?"

b. A-Ying **u** thak cit-pun che, (i) kam **u**
A-Ying perf read one-class book s/he Q perf
"A-Ying read the book, did he?"

Danish: *NEG-NEG / *POS-POS

Same-polarity tags in Danish

(99) a. *Jasper læste **ikke** bogen, gjorde han **ikke**
Jasper read.past not book.def did he not
"Jasper didn't read the book, didn't he?"

b. *Jasper **læste** bogen, **gjorde** han
Jasper read.past book.def did he
"Jasper read the book, did he?"

Samoan: *NEG-NEG
(POS-POS status currently unknown)

Same-polarity tag in Samoan

(100) na **le'i** **faitau** e Sina le tusi, (na) **le'i** **faitau**-a
past neg read erg Sina abs.the book past neg read-ES
"Sina didn't read the book, didn't she?"

Scottish Gaelic: *NEG-NEG / POS-POS

Same-polarity tags in Scottish Gaelic

(101) a. ***cha** do leugh Calum an leabhar sin, **nach** do **leugh**
neg past read Calum the book that neg.Q past read
"Calum didn't read that book, didn't he?"

b. **leugh** Calum an leabhar sin, an do **leugh**
read Calum the book that Q past read
"Calum read that book, did he?"

Persian: *NEG-NEG / *POS-POS

Same-polarity tags in Persian

(102) a. *Naysan ketaab-o **na-khoond**, **na-khoond**
Naysan book-obj neg-read neg-read
"Naysan didn't read the book, didn't he?"

b. *Naysan ketaab-o **khoond**, **khoond**
Naysan book-obj read read
"Naysan read the book, did he?"

Brazilian Portuguese: *NEG-NEG / POS-POS

Same-polarity tags in Brazilian Portuguese

(103) a. *o Bruno **nao** **leu** o livro, **nao** **leu**
the Bruno neg read the book neg read
"Bruno didn't read the book, didn't he?"

b. o Bruno **leu** o livro, **leu**
the bruno read the book read
"Bruno read the book, did he?"