

# English Tag Questions: Corpus Findings and Theoretical Implications\*

Jong-Bok Kim and Ji-Young Ann  
Kyung Hee University, Seoul  
jongbok@khu.ac.kr, lianchanster@gmail.com

May 25, 2008

## Abstract

In terms of formal properties, canonical English tag questions are sensitive to three main factors: the choice of auxiliary and pronoun, polarity (negation), and intonation pattern. Even though the general uses of tag questions follows the described constraints, their actual use in real life appears to be much more complex. This paper aims to report the corpus findings of English tag questions from the ICE-GB (International Corpus of English, Great Britain) and show that the corpus data reveal complex variations. In particular, we discuss the properties of reverse and constant polarity tag, situational tag, subjectless tag constructions. We then sketch a constructional analysis that can capture the fact that canonical as well as noncanonical tag question constructions all have much in common, but differ among themselves. We also hint that all these tag constructions are linked as a network of constructions in which specific constructions inherit general properties from their supertype constructions while they have their own constructional constraints.

Key words: English tag question, auxiliary, polarity, pronoun, ICE-GB, constructions

---

\*An earlier version of this paper was presented at the 2008 Northwest Linguistics Conference held at University of Washington at Seattle, May 3–4, 2008. We thank the participants of the conference for their questions and criticism. We also thank anonymous reviewers of this journal for helpful suggestions. All errors are of course ours.

# 1 Introduction

English tag questions, added to the main statement of a sentence, have various usages as noted by Tottie and Hoffmann (2006) and others:

- (1) Informational:  
A: You're getting paid for this, are you?  
B: Twenty five quid.
- (2) Confirmatory:  
A: I am gonna try to go walking for a little bit. I don't need a jacket, do I?  
B: No, It's still pleasant.
- (3) Attitudinal:  
A: She'll be in trouble, won't she?  
B: mh...
- (4) Facilitating  
A: Right, it's two, isn't it?  
B: Mm.
- (5) Challenging:  
A: You put what?  
B: Put six eggs on, didn't I? anyhow, I am putting two on.

As illustrated here, tag questions can be used as (a) asking information, (b) expecting confirmation from the hearer, (c) emphasizing what the speaker says, (d) making sure of the truth of what the speaker says to facilitate the conversation, (e) challenging a statement.<sup>1</sup>

In terms of formal properties, as is well known, the formation of canonical tag questions is sensitive to at least three factors such as the auxiliary, pronoun, and polarity value of the main sentence that the tag question is attached to:<sup>2</sup>

---

<sup>1</sup>Tottie and Hoffmann (2006) classifies 'challenging' into two subtypes such as 'peremptory' and 'aggressive'.

<sup>2</sup>The intonation of the tag question also determines its usage:

- (6) a. It would probably work in the corner, wouldn't it?/\*wasn't it?
- b. Well it sounds quite good, doesn't it?/\*doesn't there?
- c. You couldn't give us a hand, could you?/\*couldn't you?

As illustrated here, the auxiliary in the tag agrees with the tense, aspect, and modality of the auxiliary verb in the anchoring clause. The polarity value of the preceding anchoring clause also affects that of the tag question: when the main clause is positive, the tag question is negative, and vice versa. The pronoun in the tag also agrees with the person, number, and gender value of the main clause's subject (cf. Catell 1973, Huddleston 1970, Tottie and Hoffmann 2006).

Even though most of the tag questions follows the general rules described here, our corpus search reveals that the real-life use allows much more variations. In the paper, we report the corpus findings of English tag questions and discuss their formal and pragmatic properties. We then briefly sketch a formal analysis that can account for such variations in tag questions.

## 2 Sources and Data Extraction

For our corpus research, we used the ICE-GB (the British component of the International Corpus of English). The corpus ICE-GB consists of about one million words of spoken and written English, a collection of 200 written and 300 spoken texts. Every text is grammatically annotated, permitting complex and detailed searches across the whole corpus. The corpus contains total 88,365 parse trees (text units), among which 59,640 are spoken texts. The corpus is accompanied by the program called ICECUP that allows us to perform a variety of different queries. We used the so-called Fuzzy Tree Fragments, in particular, the function and category value set 'tag question' and 'clause'. Using this method, we found total 754 tag questions from the corpus.<sup>3</sup>

- 
- (i) a. He doesn't know what he's doing, does he? (falling)
  - b. You are coming, aren't you? (rising)

The rising intonation is used when soliciting information whereas the falling one is to strengthen the statement or seeks the agreement of the hearer.

<sup>3</sup>We excluded the following three examples the corpus marks as tag questions:

- (i) a. Twenty-third, is that?'<S1A-030 107>
- b. And you suffer from mild asthma, is that right?' <S1A-051 106>

As shown in the table in (7), of these 754 instances, we have only 15 tag questions in the written texts and the remaining 740 instances in the spoken texts, supporting the general assumption that tag questions are predominant in spoken language:

(7) Frequency of TQ (tag questions) in the ICE-GB

text type			# of text units	TQ Frequency
spoken	dialogue	private	29993	470 (62%)
		public	14717	241 (32%)
	monologue		16191	8 (1%)
written	non-printed		7234	14 (2%)
	printed		20230	21 (3%)
	total		88365	754

The table tells us that private conversations (62% of the total tag questions) are the most preferred text types for tag questions that include text types such as classroom lessons, broadcast discussions, broadcast interviews, debates, and business transactions. This preference is due to the general tendency that tag questions are most common in colloquial usages.

### 3 Formal Properties of Tag Questions

#### 3.1 Polarity Types

As noted in the beginning, tag questions are at first sensitive to polarity value. The general rule is that the polarity value of the tag question is the reverse of the polarity value of the main clause that the tag question anchors to. About 90% of the total tag questions (663 of total 754) follows this general rule in forming tag questions:<sup>4</sup>

(8) Positive-Negative

- 
- c. That's a fantastic investment each year in that and in my belief and one watches these things, doesn't one, as one goes about it pays off' <S2A-031080>

We assume that the pronoun 'that' and 'one' cannot function as a pronoun in the tag question.

<sup>4</sup>The notation 'S' means a spoken text and the remaining numbers indicate the text identification number. To increase the readability of the corpus data, we also added brackets to the tag questions, or edited punctuation marks and spoken style expressions.

- a. You've seen Martin's hall, [haven't you]? <S1A-073-004:1:B>
  - b. Now they meet on the on the packet sailing from England, [don't they]? <S1 A-020-207:1:B>
- (9) Negative-Positive
- a. Nigel you couldn't give us a hand, [could you]? <S1B-074-213:3:A>
  - b. Oh I didn't pay the Woodville Feast by chit, [did I]? <S1A-039-247:1:B>

Unlike such reverse polarity examples, we also have cases where the polarity value of the tag question is identical with that of the anchoring clause (cf. cattell 1973, Culicover 1992, Holmes 1982):

- (10) Positive-Positive
- a. But the wall will stand, [will it]? <S1A-009-204:1:A>
  - b. You went on that day, [did you]? <S1A-029-224:1:B>
  - c. The Dickens itself reads more like Dickens, [does it]? <S1B-026-136:1:A>
- (11) Negative-Negative
- Well you aren't supposed to record the systolic when you can hear more than one sound occurring you know successively, [aren't you] <S1B-004-150:1:A>

As noted by Kay (2002) and Kimps (2007) and others, the constant polarity tags in general modalize the proposition in the host clause or solicit positive responses. In other words, the constant polarity tag is attached to a sentence that the speaker is not putting forward as his own but is citing in order to ask the listener if it is his. It also functions as the turn-allocation, conducive to a positive response.

The frequency of these four types can be summarized as following:

- (12) Frequency of Tag Questions by Polarity Values:

Types	Pos-Neg	Neg-Pos	Pos-Pos	Neg-Neg	Total
Frequency	467 (62.5%)	196 (26%)	88 (12%)	3(0.5%)	754

An interesting set of data with respect to the polarity is examples where *not* is not contracted, occurring right after the personal pronoun:

- (13) a. Well now things in a hotel may break from time to time [may they not]? <S1B-067-47:1:A>
- b. At that point, Mr Hook, on behalf of Ferndale had indicated to you that he was interested in further expansion of his business, [had he not]?
- c. A calculation about morality really is an incompatibility, [is it not]? <S1B-060-082:1:A>

As such, non-contracted tags, though a little unusual, can be found in the corpus.

### 3.2 Auxiliary Agreement

Various different types of auxiliary verbs occur in tag questions. The choice of the auxiliary verb in tag questions depends on the modality, tense, and aspect of the main sentence's auxiliary verb:

- (14) a. But it was an important point, [wasn't it] <S1B-063-276:1:B>
- b. You turned it round then, [didn't you]? <S1A-083-018:1:B>
- c. And we've got your note on page two eight seven , [haven't we]? <S1B-069-194:1:A>
- d. Oh I could try that then, [couldn't I] <S1A-017-253:1:B>

The following table shows the frequency of tag questions by the type of auxiliary verbs:

- (15) Frequency of the TQ by the type of auxiliary verb

Aux	be	do	have	modal	Total
Frequency	404 (54%)	203 (27%)	51 (7%)	96 (13%)	754

As noted in the table, the auxiliary *be* and its inflected forms are the most frequently used auxiliary, followed by *do*. This frequency tendency is similar to what Tottie and Hoffmann (2006) found from the BNC corpus.

If we look at the usage of modal verbs, the modal *will* (including *would*) has the highest frequency, followed by the modal *can*:

(16) Frequency of the Tag question by the type of modals

Modals	will	can	shall	may	must	Total
Frequency	50	32	7	5	2	96

The high frequency of using *be*, *do*, and *will* seems to be related to the discourse function of tag questions. As noted by Tottie and Hoffmann (2006), the confirmatory, facilitating, and attitudinal use of tag questions occupies more 90% of the total use. These three auxiliary verbs seem to easily match these functions.

As noted earlier, the auxiliary in the tag question canonically agrees with that of the auxiliary tense, modality, and aspect. However, the corpus also reveals total 16 instances (2%) where the auxiliary in tag disagrees with that of the anchoring clause. A simple disagreement is related to tense information:

- (17) a. So it didn't sound too good, [does it]? <S1A-094-132:1:B>  
b. Oh that'll be handy, [wouldn't it]? <S1A-019-241:1:A>

We even have cases where the auxiliary in tag questions is not linked to the auxiliary in the main clause:

- (18) a. It'd be about a half-hour journey, [isn't it]? <S1A-019-243:1:C>  
b. We'll interrupt it at that point to just do a quick statistical analysis of all your data, [shall we]? <S1B-004-002:1:A>  
c. ...We ought to be doing the same thing within the faculty, [should we]?...<S1B-075-082:1:E>

The data indicate that the auxiliary in the tag question is sometimes chosen, depending on the illocutionary force the speaker intends to express.

### 3.3 Pronoun agreement

The tag question, consisting of an auxiliary and a pronoun, also matches with the pronoun value of the anchoring sentence, as also observed from the corpus:

- (19) a. I am right in that, [aren't I]? <S1A-028-002:1:A>  
b. You made her bottle out, [didn't you]? <S1A-070-204:1:B>

- c. But we started the week before you though, [didn't we]? <S1A-040-179:1:C>

The proportions of pronouns in the tag questions we found can be summarized as following:

- (20) Distribution of Tag questions by the pronoun type

	Pronoun Type	Frequency
personal	1st per (I, we)	53 (7%)
	2nd per (you)	172 (23%)
	3rd per (he, she)	69 (9%)
	3rd per pl (they)	73 (10%)
nonpersonal	it	356 (47%)
	there	31 (4%)
total		754

As shown here in the table, the most frequently used pronoun in tag questions is the nonpersonal third-singular *it* followed by the 2nd person pronoun *you*. The high frequency of *you* is expected from the predominant usage of tag questions in dialogue such as direct conversation, classroom lesson or discussion. The pronoun can be other than personal pronouns:

- (21) a. It was an important point, [wasn't it] <S1B-063-276:1:B>  
 b. There's no problem, [is there]? <S1B-013-023:1:A>

The pronoun *there* has the lowest frequency among the other pronouns. The pronoun *there* appears only with the *be* auxiliary in the question:<sup>5</sup>

- (22) a. There's not much point in that, [is there]? <S1A-038-223:1:A>  
 b. There wouldn't be any point, [would there]? <S1A-029-190:1:D>

The ICE-GB corpus again gave us more complex data in the pronoun agreement. For example, the agreeing pronoun, as well as the auxiliary verb, need not be the subject of the matrix clause. It can be the embedded subject:

- (23) a. I mean that's not on, [is it]? <S1A-007-079:1:B>

<sup>5</sup>Except the one in (22)b, *there* in the tag occurs only with the auxiliary *be*.



- b. I think that it was Sunday, [wasn't it] <S1A-023-125:1:A>

A complexity arises from examples like (24) where we at first glance see no clear agreement between the pronoun in the tag question and the putative anchoring pronoun in the main clause:

- (24) a. But everybody talks about them you see, [don't they], as being so marvelous? <S1A-016-203:1:D>  
b. Everyone does at least one dish, [don't they]? <S1A-071-360:1:D>  
c. Everybody made it one three two or above, [did they]? <S1B-004-080:1:A>

As noted by Bender and Flickinger (1999), these uses indicate that the pronoun *they* can be used as referring to a single unit.

We further observe in the so-called 'situational tags' in which the pronoun in tag question has no related anchoring pronoun in the preceding main or embedded clause:

- (25) a. I bet there's a load of old rubbish put on these forms, [don't you]? <S1A-007-251:1:B>  
b. She looks Puerto Rican or something, [is it]? <S1A-058-008:1:C>

What we have seen so far is that even though most of the tag questions we found from the corpus observe the general formation rules with respect to the use of auxiliary, polarity, and pronoun, there also exist unexpected, numerous variations that override these rules. In what follows, we observe further variations we found from the corpus.

## 4 Variations in Tag Questions

### 4.1 Tag Questions in the Embedded Sentence

Though the tag question is canonically linked to the matrix clause, the corpus data reveal the higher proportions of the tag questions linked to the embedded sentence (cf. Bender and Flickinger 1999). As noted in the literature, the tag questions attached to the embedded clause are generally used when the main clause has the so called 'hedging verb' which weakens or softens the speaker's assertiveness about the proposition uttered. The hedging verbs searched in ICE-GB are such as *suppose*, *mean*, *see*, *say*, *think*, *know*, etc:

- (26) a. I suppose it would be optical scanning, [would it]? <S1A-024-041:1:A>
- b. I mean it it sounds a bit holy, [doesn't it]? <S1A-003-120:1:B>
- c. You know there are few people like that in whatever field, [aren't there]? <S1A-045-008:1:B>

Another interesting type of tag questions attached to the embedded question includes constructions like ‘*the pity is that..., that’s what..., this is why..., one thing is that...*’:

- (27) a. The pity is often they’re discouraged from gardening, [aren’t they] because they don’t know where to start <S1B-025-113:1:A>
- b. That’s what it was, [wasn’t it] really <S1A-052-037:1:B>
- c. One thing that I find interesting is that that the dramatists both tragedy the tragedians and the comic playwrights tend to use the framework of the single day, [don’t they] <S1B-019-007:1:A>

When considering that the subject here functions as a discourse marker, we can expect that the tag is anchored to the predicative sentential complement.

We also find tag questions in adverbial subordinate clauses with conjunctions like *when*, *as*, or *because*:

- (28) a. Well I won’t because it will weigh extra, [won’t it]? <S1A-011-222:2:A>
- b. It’s surely quite difficult these days to persuade an actor or actress to commit themselves for what six to eight months when there is always the possibility, [isn’t there] of lucrative television work <S1B-050-072:1:A>

The data indicate that tag questions can be attached to any propositional sentence to perform the desired illocutionary force such as confirmation and challenging.

## 4.2 Subjectless Tagged Sentence

Of total 754 tag questions we found from the corpus, 53 (about 7%) examples have missing subjects. These subjectless tagged sentences (STS) can be basically classified into four types as noted by Kay (2002):

- (29) a. Type I: a tensed main verb in the host sentence and support *do* in the tag  
 b. Type II: a tensed auxiliary which is repeated in the tag  
 c. Type III: a nonverbal predicate and finite *be* occurs in the tag  
 d. Type IV: a non-finite verb in the host that is an appropriate complement for the tag auxiliary

The ICE-GB corpus give us all these four types:

- (30) a. Sounds like the masons really, [doesn't it] <S1A-027 #058:1:C>  
 b. Would be a bit of a bad deal, [wouldn't it] <S1A-030-033:1:B>  
 c. Confirmation in this hand, [isn't there] <S1A-067-323:1:B>  
 d. Coming back from work, [was she] <S1A-020-228:1:B>

Of these four types, Type I and III are more frequently used as shown in the table:

- (31) Frequency of STS Types

Type	Frequency
Type I	11 (20%)
Type II	8 (15%)
Type III	29 (54%)
Type IV	5 (1%)
Total	53

Even if the subject or the auxiliary is unrealized in the STS constructions, the subject and auxiliary in tag question still agrees with the unrealized subject or auxiliary:

- (32) a. Looks as if it might well have been, [doesn't it] <S1A-023-286:1:A>  
 b. Fascinating, [isn't it] <S1B-032-139:1:A>  
 c. Could suggest loss of virginity [couldn't it] <S1B-014-015:1:A>  
 d. Can't fit me in as well, [can you] <S1A-039-361:1:A>

This means STS constructions are not erratic constructions but follow the general rules in forming English tag questions.

### 4.2.1 Tag Questions in Nondeclaratives

Tag questions are also used in imperatives, with adding *will you* or *won't you?*

- (33) a. Don't look back at that, will you? <S1A-088-224:1:A>
- b. Excuse me, won't you? <W1B-003-160:2>
- c. Don't tell, will you? <S1A-032-182:1:A>
- d. Let's stop for the moment, shall we? <S1A-001-051:1:A>

Literature points out that tag questions can be attached to exclamative sentences as in (34), but we found no instance from the corpus:

- (34) What a pretty dress that is, isn't it? (Hudson 1975)

One interesting fact is that unlike the traditional assumption, we found a few cases where the tag question is anchored to an interrogative sentence:

- (35) a. Do you need to be qualified, don't you, teaching English? <S1A-035-029:1:B>
- b. Didn't Mr Hook say to you that at this point that he was interested in further expansion of the Ferndale business, did he not? <S1B-064-118:1:A>

Considering the discourse function that tag questions carry, there seems to be no clear reason to rule out interrogatives as long as the intended illocutionary force can be achieved.

### 4.3 Situational Tagging

There are also examples where neither the auxiliary verb nor the pronoun in the tag question agrees with their targets in the main statement.

- (36) a. What you want is a little greenhouse really, [don't you]? <S1A-007-27:1:B>
- b. I think he's very sensible, [don't you] <S1A-011-214:2:A>
- (37) a. The person that's affected is me basically [isn't it] <S1A-026-75:1:A>

- b. So you put in the sedative [is it] and uh <S1A-089-77:1:B>
- (38) a. She's planning a wild party, [aren't you]? <S1A-019-351:1:F>
- b. Brandon's going to lend me loads of money, [aren't you] cos my cash and money are stolen <S1A-040-368:1:D>

Most of the subject in the situational tag construction is the 2nd singular pronoun *you* or the 3rd singular pronoun *it*, as shown in the following frequency table:

- (39) Aux Type in Situational Tags:

Aux Type	Frequency
isn't/is/was/wasn't it	14 (51%)
do/don't/did/didn't you	11 (40%)
aren't you	2 (9%)
total	27

In terms of the function of the situational tag in the corpus, we could observe that the main function of such situational tags is to add questioning and confirming force with some pausing time, similar to the so called invariable tag questions with expressions like *huh?* or *right?*:

- (40) And now we're s we're sitting in a in a house on a piece of ground where you had kept ponies, isn't it <ICE-GB:S1A-028-184:1:B>

The next frequent uses are related to hedging verbs, as observed from the following conversation:

- (41) B: I mean they surely they don't keep hold of everybody <ICE-GB:S1A-007-288:1:B>

A: I don't believe it, do you? <ICE-GB:S1A-007-289:1:B>

As noted in the conversation, the tag part *do you* intended to ask the question *do you believe it?*, seeking the hearer's opinion.

#### 4.4 Variations in the Position

We have already seen that tag questions can be linked to various types of embedded sentences as well as adverbial subordinate sentences. Bender and Flickinger (1999) assume that the tag phrase must modify the full sentence, attaching at S, based on the fact that the tag question appears outside of elements extraposed from subject NPs:

- (42) a. A review appeared of Chomsky's latest book, didn't it?  
b. That movie bombed that you really liked, didn't it?

However, we can observe that the tag phrase can appear before the extraposed expression:

- (43) a. It is marvelous, [isn't it], to see cabinet ministers shadow cabinet people uh very distinguished journalists all these people turning up in a place like Neath where a few years ago people said<S1B-029-022:1:A>  
b. That's very expensive, [isn't it], doing it<S1A-035-039:1>

In addition to these, the real corpus data reveal striking variations in the positional distribution of the tag question. The tag question can be attached even to a VP projection:

- (44) a. You don't know, [do you], whether the hotel had any system at all for checking the fittings round the swimming pool <S1B-067-072:1:A>  
b. You don't, [do you], want us to think of faith as a synonym for tradition <S1B-028-103:1:C>

The tag question can be even occur within an NP structure too:

- (45) a. Uhm in fact we know that there was a rupture, [don't we], of the ligament <S1B-068-141:2:D>  
b. It 's a funny thought [isn't it] that I was embarrassed<S1A-032-037:1:A>  
c. and of course it has a connection, [hasn't it], with University College<S1A -033-022:1:A>

It can also occur before an adjectival complement:

- (46) a. These are quite smart, [aren't they], for the schools conference  
super badge<S1A-077-357:1:D>
- b. But Paddy it would be compatible, [would it not], with the  
U N resolutions for Saddam to remain in power ...<S1B-027 -  
040:1:A>

The observations indicate that the positional possibilities of tag questions are much more flexible than the literature has assumed.

## 5 Theoretical Implications

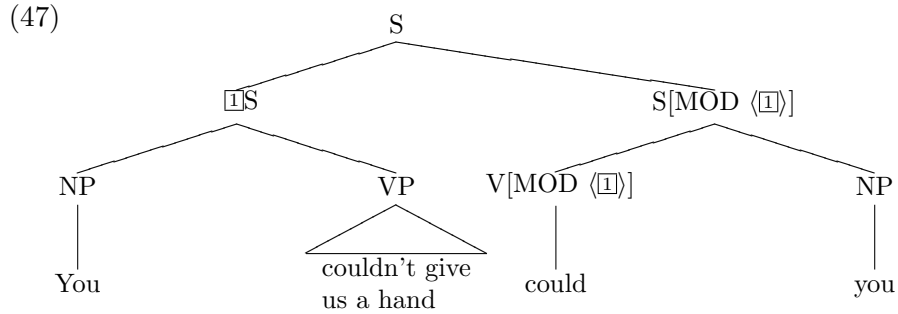
As we have seen so far, the corpus search supports the general assumption that English tag questions are dominant in spoken texts. At the same time, we have also noticed that English tag questions display much variations, depending on values such as polarity, auxiliary, pronoun, illocutionary force, and so forth. Given the complexity of tag questions, we believe that a suitable grammar needs to

- refer to not just syntactic properties, but also other grammatical properties such as semantic and pragmatic ones
- allow the tight interfaces among such properties
- represent the constructional properties of canonical as well as non-canonical tag questions

The grammar we adopt that appears to observe these is HPSG (Head-Driven Phrase Structure Grammar). Let's us briefly see how tag questions can be accounted for within HPSG (cf. Sag et al. 2002, Sag 2007).

### 5.1 Syntactic Structure

The first question of tag questions concerns their syntactic structure. Following the previous literature (cf. Bender and Flickinger 1999), at first glance we assume that tag questions are modifiers to an S as represented in the tree structure:



To be more precise and more formal, we can assume that the MOD information is originated from the auxiliary in the tag in accordance with the following lexical construction (cf. Bender and Flickinger 1999):

(48) Tag Aux Lexical Construction:

$$\left[ \begin{array}{l} \text{HEAD | AUX +} \\ \text{MOD } \langle \{ \textit{verbal} \} \rangle \\ \text{SUBJ } \langle \quad \rangle \\ \text{COMPS } \left\langle \left[ \begin{array}{l} \textit{pronoun} \\ \text{IND } \mathbb{I} \end{array} \right] \right\rangle \\ \text{RELS } \langle \quad \rangle \\ \text{FORCE } \textit{illoc-force} \end{array} \right] \rightarrow \left[ \begin{array}{l} \text{HEAD | AUX +} \\ \text{SUBJ } \langle \text{XP}[\text{IND } \mathbb{I}] \rangle \\ \text{RELS } \textit{nelist} \end{array} \right]$$

This lexical construction means an auxiliary verb (in the right daughter) can be projected into a tag auxiliary in the left mother. The mother tag auxiliary can modify a *verbal* expression and selects as its complement a personal pronoun whose IND value is identical with the input auxiliary's subject.<sup>6</sup> In addition, this lexical process ensures that the auxiliary verb has no semantic content to contribute (indicated in the RELS value) other than illocutionary force. For example, consider the canonical use of *could* in the tag:<sup>7</sup>

<sup>6</sup>We thus allow a tag question to modify any verbal expression including VP or S.

<sup>7</sup>As the types of *illoc-force* (illocutionary force), we assume *asserting*, *promising*, *communicating*, *exclaiming*, *inquiring*, and *ordering*.

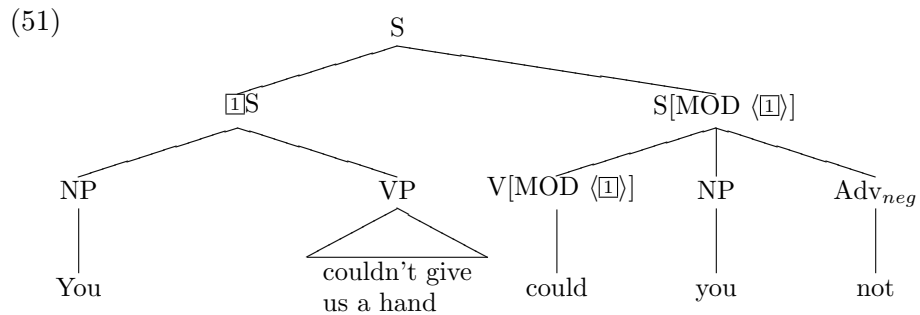


$$(49) \left[ \begin{array}{l} \langle \text{could} \rangle \\ \text{HEAD} \mid \text{AUX} + \\ \text{MOD} \langle [\textit{verbal}] \rangle \\ \text{SUBJ} \langle \quad \rangle \\ \text{COMPS} \langle \text{NP}_i[\textit{pro}] \rangle \\ \text{RELS} \langle \quad \rangle \\ \text{FORCE} \textit{asserting} \end{array} \right] \rightarrow \left[ \begin{array}{l} \langle \text{could} \rangle \\ \text{HEAD} \mid \text{AUX} + \\ \text{SUBJ} \langle \text{NP}_i \rangle \\ \text{COMPS} \langle \text{VP}[\text{IND } s1] \rangle \\ \text{RELS} \langle \left[ \begin{array}{l} \text{PRED } \textit{could-rel} \\ \text{ARG1 } s1 \end{array} \right] \rangle \end{array} \right]$$

As seen from the right daughter, the canonical *could* selects one NP subject and a VP complement with the meaning of *could-rel*. This one can be realized as a tag *could* in the left mother, adding the illocutionary force with no semantics. Notice that, given the analysis of Kim and Sag (2002) in which the negator *not* is introduced as the complement of a finite auxiliary verb, the output mother lexical construction will be slightly different:

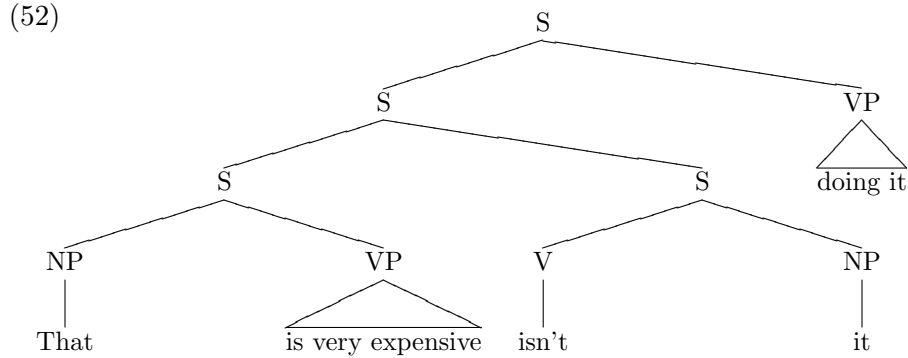
$$(50) \left[ \begin{array}{l} \langle \text{could} \rangle \\ \text{HEAD} \mid \text{AUX} + \\ \text{MOD} \langle \text{S} \rangle \\ \text{SUBJ} \langle \quad \rangle \\ \text{COMPS} \langle \text{NP} \left[ \begin{array}{l} \textit{pro} \\ \text{IND } \boxed{1} \end{array} \right], \text{Adv}_{neg} \rangle \\ \text{RELS} \langle \quad \rangle \\ \text{FORCE} \textit{asserting} \end{array} \right] \rightarrow \left[ \begin{array}{l} \langle \text{could} \rangle \\ \text{HEAD} \mid \text{AUX} + \\ \text{SUBJ} \langle \text{NP}[\text{IND } \boxed{1}] \rangle \\ \text{COMPS} \langle \text{Adv}_{neg}, \text{VP}[\text{IND } s1] \rangle \\ \text{RELS} \langle \left[ \begin{array}{l} \text{PRED } \textit{could-rel} \\ \text{ARG1 } s1 \end{array} \right] \rangle \end{array} \right]$$

The negative *could* now can select an adverbial negator and a VP as its complement. This output in the mother can be realized as a Tag Aux lexical construction, eventually allowing us to generate a structure like the following in our analysis:



As noted in (13), the corpus give us examples where the negator is not contracted but appears after the subject.

The present analysis in which the tag functions as a modifier can also be applied to positional variations. Given that the extraposition is an attachment to an S, we will have a structure like the following:<sup>8</sup>



## 5.2 Constructional Constraints and Grammatical Interfaces

As noted so far, tag questions closely interact with syntactic, semantic, and pragmatic information. As we have seen, the canonical tag questions are sensitive to the auxiliary, pronoun, and polarity value. Given that the tag syntactically functions as a modifier, the issues remain of how to make sure the auxiliary, subject, and polarity value of the tag part refers to the counterparts in the anchoring clause. Such constraints are peculiar to the tag question constructions, leading us to posit a constructional rule like the following (cf. Kim and Sells 2008):

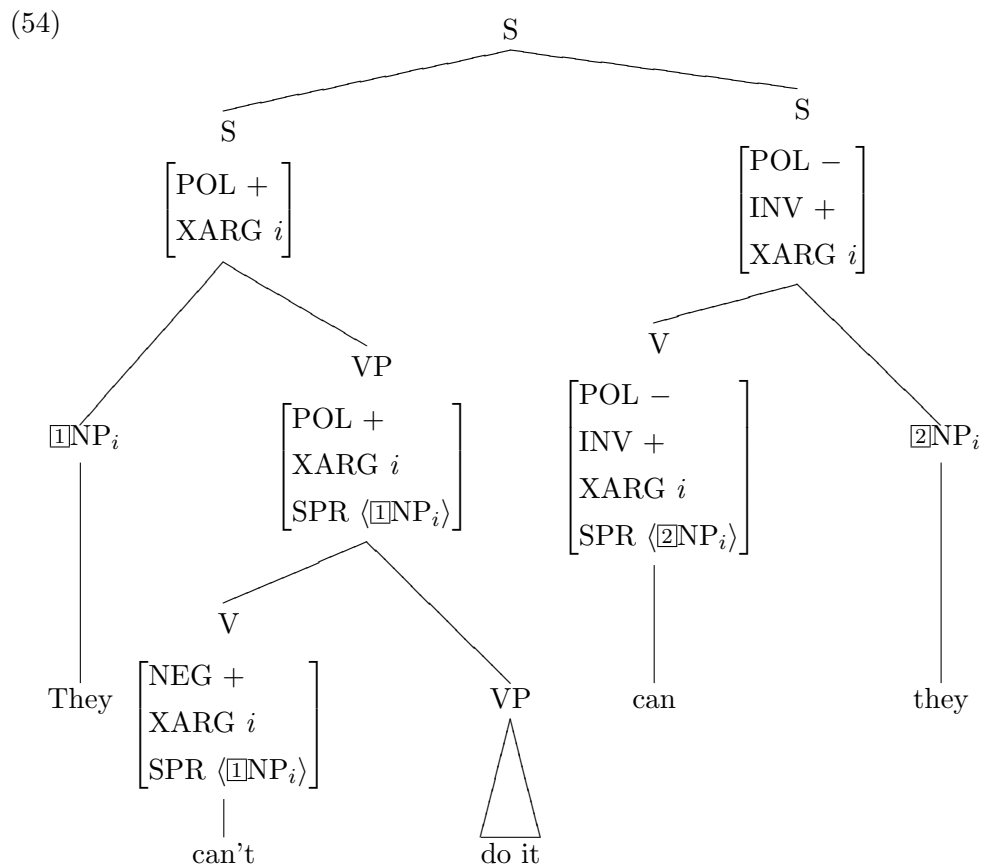
(53) Canonical Tag-Question Construction:

$$S \rightarrow S \begin{bmatrix} \text{POL } \boxed{1} \\ \text{AUXFORM } \boxed{2} \\ \text{XARG } i \end{bmatrix}, S \begin{bmatrix} \text{MOD } \langle \boxed{4} \rangle \\ \text{POL } \neg \boxed{1} \\ \text{INV } + \\ \text{AUXFORM } \boxed{2} \\ \text{XARG } i \\ \text{FORCE } \textit{illoc-force} \end{bmatrix}$$

<sup>8</sup>We assume that the extraposition can occur either before or after the tag question, contra to Bender and Flickinger (1999).

This constructional rule in (53), reminiscent to those given by Kay (2002), ensures several constructional constraints: the tag part functions as a modifier to the anchoring sentence; it has an reverse POL(ARITY) value to the anchoring clause's POL value; it has subject-auxiliary-inversion; its AUXFORM is identical with the anchoring clause's AUXFORM; its subject index represented by the feature XARG is linked to the anchoring clause's XARG value; the tag part further is elided.<sup>9</sup>

This constructional rule will then project a structure like the following:



<sup>9</sup>Traditionally, arguments are classified into external and internal ones in which the former usually refers to the subject. The introduction of such a semantic feature is necessary if we want to make the subject value visible on the S node (see Bender and Flickinger 1999). That is to say, although a VP has a SPR value for its subject, once the VP and subject combine, the resulting S no longer has any information about any features of the subject – including its semantic index. XARG is a mechanism for making this information visible at the S level, which is where the tag question adjoins.

As represented here, the POL feature of the matrix verb is passed up to the first S. The tag question then needs to have the opposite POL value in accordance with the rule in (53). The semantic feature XARG identified with the subject starts from the auxiliary verb and then is semantically composed into the meaning of VP and then S. The XARG value in a sense makes the subject’s index value visible at the top level of the sentence in question so that the tag subject can also refer to this.

The Tag Question Construction Rule in (53) will just allow canonical reverse tag questions. As we have noticed, however, there exist noncanonical tag questions like constant polarity tag, situational tag, and STS (subjectless tag sentence) constructions. As we have seen so far, these noncanonical constructions have their own syntactic, semantic, and constructional constraints while inheriting some basic properties of tag questions such as the subject-auxiliary inversion and VP ellipsis, as noted by Kay (2002). For example, the STS will have the following constructional constraint:

(55) Subjectless Tag Sentence Construction:

$$S \rightarrow S_{\boxed{4}} \left[ \begin{array}{l} \text{SUBJ } \langle XP_i \rangle \\ \text{POL } \boxed{1} \\ \text{AUXFORM } \boxed{2} \\ \text{XARG } i \end{array} \right], \quad S \left[ \begin{array}{l} \text{MOD } \langle \boxed{4} \rangle \\ \text{POL } \neg\boxed{1} \\ \text{INV } + \\ \text{AUXFORM } \boxed{2} \\ \text{XARG } i \\ \text{FORCE } \textit{illoc-force} \end{array} \right]$$

As noted here, the only difference from the general rule in (53) is that the anchoring sentence’s subject is a *pro*.<sup>10</sup> All the other constraints remain intact.

Unlike these two main types of tag questions, the situational tag does not refer to the main clause’s subject, auxiliary, or polarity value. The only thing it refers to is the semantics of the main clause:

(56) Situational Tag Sentence Construction:

$$S \rightarrow S_{\boxed{4}} \quad S \left[ \begin{array}{l} \text{MOD } \langle \boxed{4} \rangle \\ \text{INV } + \\ \text{FORCE } \textit{asserting} \end{array} \right]$$

---

<sup>10</sup>See Sag et al. (2002) for the treatment of the imperative clause in which the subject is not realized.

Though we cannot do justice to how the FORCE value is realized in situational tags licensed by a limited number of auxiliary and pronoun, the situational tag functions as a confirming force for the main clause.

The scope of this paper limits our attempt to provide a detailed account for all the types of tag question constructions, but we can assume that canonical as well as noncanonical tag question constructions all have much in common, but differ among themselves, indicating that they are all linked as a network of constructions in which specific constructions inherit general properties from their supertype constructions while they have their own constructional constraints.

## 6 Conclusion

Of the total 754 tag questions we found from the ICE-GB, about 90% of the examples follow the general rules in forming tag questions. This means that more than 10% override these general rules to achieve various illocutionary force.

These non-canonical examples also display the properties found from canonical tag constructions, hinting that there exist a family of tag constructions linked as a network. Though there remain more issues in providing a precise account and figuring out the functional aspects of the tag question constructions, we have observed that both canonical and noncanonical tag questions are basic aspects of the English grammar.

## References

- Bender, Emily. and Flickinger, Daniel. P. 1999. Peripheral Constructions and Core Phenomena: Agreement in Tag Questions. In Webelhuth Gert, Jean-Pierre Koenig and Andreas Kathol(eds.), *Lexical and Constructional Aspects of Linguistic Explanation*, 199-214, Stanford: CSLI Publications.
- Cattell, Ray. 1973. Negative Transportation and Tag Questions. *Language* 49: 612-639
- Culicover, Peter. 1992. English Tag Questions in Universal Grammar. *Lingua* 88: 193-226
- Holmes, Janet. 1982. The Functions of Tag Questions. *English Language Research Journal* 3: 45-65
- Hudson, Richard A. 1975. The Meaning of Questions. *Language* 51: 1-31

- Huddleston, Rodney. 1970. Two Approaches to the Analysis of Tags. *Journal of Linguistics* 6: 215-222
- Kay, Paul. 2002a. English Subjectless Tagged Sentences. *Language* 78(3): 453-481
- Kim, Jong-Bok, and Ivan A. Sag. 2002. Negation without Head Movement. *Natural Language and Linguistic Theory* 20.2: 339-412.
- Kim, Jong-Bok and Sell, Peter. 2008. *English Syntax: An Introduction*. CSLI Publications, Stanford CA.
- Kimps, Dittie. 2007. Declarative Constant Polarity Tag Questions: A Data-driven Analysis of Their Form, Meaning and Attitudinal Uses. *Journal of Pragmatics* 39(2): 270-291
- Sag, Ivan. A, Tomas Wasow, and Emily Bender. 2002. *Syntactic Theory: A Formal introduction*. CSLI publications, Stanford CA.
- Sag, Ivan. A. 2007. Sign-Based Construction Grammar: An Informal Synopsis. Unpublished ms, Stanford University
- Tottie, Gunnel and Hoffmann, Sebastian. 2006. Tag questions in British and American English. *Journal of English Linguistics* 34(4): 283-311.