

The prosody of question tags in English¹

NICOLE DEHÉ and BETTINA BRAUN

Universität Konstanz

(Received 24 June 2012; revised 12 September 2012)

The prosodic realization of English question tags (QTs) has received some interest in the literature; yet corpus studies on the factors affecting their phrasing and intonational realization are very rare or limited to a certain aspect. This article presents a quantitative corpus study of 370 QTs from the *International Corpus of English* that were annotated for prosodic phrasing and intonational realization of the QT and the host. Factors tested were polarity, position in the sentence and the turn as well as verb type. Generally, prosodic phrasing and intonational realization were highly correlated: separate QTs were mostly realized with a falling contour, while integrated QTs were mostly rising. Results from regression models showed a strong effect of polarity: QTs with an opposite polarity were more often phrased separately compared to QTs with constant polarity, but the phrasing of opposite polarity QTs was further dependent on whether the QT was negative or positive (more separate phrasing in negative QTs). Furthermore, prosodic separation was more frequent at the end of syntactic phrases and clauses compared to phrase-medial QTs. At the end of a turn, speakers realized more rising contours compared to QTs within a speaker's turn. Verb type also had an effect on the phrasing of the tag. Taken together, our results confirm some of the claims previously held for QTs, while others are modified and new findings are added.

1 Introduction

Question tags (QTs; also: *tag questions*) are clausal elements such as the italicized ones in (1):

- (1) (a) John is a genius, *isn't he*, Peter
(b) Paula didn't go to the party, *did she*, last night
(c) So you missed another class, *did you*

Formally, they consist of an operator (auxiliary verb that corresponds to the verb in the preceding clause or *do*) and a pronominal subject (in that order), as well as an optional negation (*n't* or *not*); see e.g. Bald (1980), Gussenhoven (1984: 118), Quirk *et al.* (1985: 810), McCawley (1998: 501), Biber *et al.* (1999: 208), Huddleston & Pullum *et al.* (2002: 893). The pronominal subject corresponds to the subject in the preceding clause; i.e. the subject and verb in the tag repeat those in the main clause in proforms. The part preceding or surrounding the tag has been referred to as *anchor* (e.g. Huddleston & Pullum *et al.* 2002: 891; Tottie & Hoffmann 2006) or *host clause*

¹ We thank Daniela Brandt and Samuel Schweizer for support with the analyses, Joanna Müller for generating the figures, and two anonymous reviewers for valuable comments on an earlier version of the manuscript. Special thanks to Anne Wichmann and Jill House for discussion and for their intonational analysis of numerous sound files.

(e.g. Cattell 1973), among others. In the present article we will refer to it as the host (clause). QTs are often reversed polarity tags, i.e. a negative tag follows a positive clause or vice versa (e.g. (1a, b)). Further examples with reversed polarity tags, taken from the British component of the *International Corpus of English* (ICE-GB; Nelson *et al.* 2002; see section 2) are shown in (2); pos-neg refers to a positive verb in the host and a negation in the QT; neg-pos to the reverse. As shown in examples (1c) and (3), constant polarity tags (*reduplicative* tags in McCawley's 1998: 501 terminology) occur, too, but according to Huddleston & Pullum *et al.* (2002: 892), reversed polarity tags are much more frequent than constant polarity tags. If the tag has the same polarity as the host clause, this is usually positive. This tendency was confirmed for both British and American English in Tottie & Hoffmann's (2006) corpus study with 5,000 QTs taken from the *British National Corpus* and 2,300 QTs from the *Longman Spoken American Corpus*, and it is also true for the set of data investigated in the present article (see table 1 in section 2.1 below).²

(2) Reversed polarity tags (a-c: pos-neg; d-f: neg-pos)

- (a) The state of the wall uh the state in which you found the wall was the result of the work that you had directed to be done to the wall *wasn't it* (ICE-GB: s1b-069 #174)
- (b) We had a big chat about this before *didn't we* (ICE-GB: s1a-006 #172)
- (c) ... 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 the possibility *isn't there* of lucrative television work (ICE-GB: s1b-050 #72)
- (d) But the Labour Party's not going to abolish private medicine *is it* Donald (ICE-GB: s1b-039 #88)
- (e) You can't blame her for that really *can you* (ICE-GB: s1a-007 #017)
- (f) You don't know *do you* whether the hotel had any system at all for checking the fittings round the swimming pool (ICE-GB: s1b-067 #072)

(3) Constant polarity tags (a-b: pos-pos; c: neg-neg)

- (a) But then there's another audience *is there* that goes to regional playhouses a rather more stolid audience (ICE-GB: s1b-050 #87)
- (b) But you were there for other reasons then *were you* (ICE-GB: s1b-042 #080)
- (c) Well you aren't supposed to record the systolic when you you can hear more than one sound occurring you know successively *aren't you* (ICE-GB: s1b-004 #150)

Huddleston (1970) postulates an underlying paratactic construction to account for the relation between QT and host. More generally, QTs have been considered parentheticals in the syntax based on their (syntactic) form and function (e.g. Knowles 1980; Quirk *et al.* 1985: 919; Ziv 1985). In his syntactic–semantic approach to clausal parentheticals, Potts (2002: 651) sees QTs as clausal expressions and treats them on a par with other types of parentheticals, among them sentence adverbs and comment clauses (e.g. *I think*). Following Emonds (1976), he argues that they have comma intonation, i.e. they are produced with an intonation break immediately before and after. QTs are syntactically defective (or: elliptical), consisting only of a subject and an auxiliary. The

² A reviewer notes that (3c) is ungrammatical to the native ear, and that the use of the constant polarity tag may be caused by the long distance between the QT and the anchor in the host, 'analogous to concord of proximity in number agreement'.

missing VP can be reconstructed from the host clause (see e.g. Knowles 1980: 396–401 for discussion). While Ziv (1985: 189) claims that tags are formally different from other types of parentheticals in that they are ‘restricted to sentence-final position’, corpus data do not support this claim: although QTs do indeed often occur sentence-finally, it is not unusual for them to occur clause-medially or even within syntactic constituents (e.g. (2c) and (4); see also Bald 1980, Knowles 1980, Tottie & Hoffmann’s 2006 corpus study and the data analysis in section 2.1 below).

(4) Non-clause final tags

- (a) You don’t *do you* want us to think of faith as a synonym for tradition (ICE-GB: s1b-028 #103)
- (b) I mean in in that piece we’ve just heard from The Revenger’s Tragedy it’s a mixture *isn’t it* of original instruments and kind of what sound to me like modern trumpets (ICE-GB: s1b-023 #140)
- (c) And uh we’ve learnt *haven’t we* that books are all right but there is nothing like uh the help you get from society members as the team have said and doing it yourself (ICE-GB: s1b-025 #166)

Assuming that QTs are parentheticals in the syntax, the prediction is that they should prosodically behave like other types of parentheticals and form a separate prosodic domain by default (Selkirk 1984: 295; Nespor & Vogel 1986: 189–90 among many others). On the other hand, it has been shown that prosodic separation of other short and prosodically light parentheticals such as comment clauses is far from obligatory (Dehé 2009 and references therein).

Approaches specifically dealing with the prosody of QTs are varied and their results on occasion conflicting. With respect to phrasing, Bolinger (1989) notes that it is common for tags containing modals such as *will* and *can*, along with other ‘appended material’ such as reporting verbs, to be unaccented and integrated into a prosodic domain as intonational tail, i.e. as post-nuclear unstressed material. However, he also notes that ‘[i]t is difficult to generalize about tag intonation because once we look beyond the central auxiliaries *will* and *can* we find that the intonation may depend on the choice of the tag’ (Bolinger 1989: 164). Given that at least *be* and *do* occur more frequently in QTs than modals (see Tottie & Hoffmann 2006), this is an important qualification. Other authors maintain that QTs should be phrased separately in general: Knowles (1980), for instance, maintains that tags have comma intonation, i.e. that they are set off intonationally and temporally (in terms of phrasing) from preceding and following host material. For his own variety he observes that within the tag, it is always the auxiliary, never the pronoun that is accented (Knowles 1980: 393; see also Quirk *et al.* 1985: 810). Selkirk (1984: 295) includes QTs with other types of parentheticals, which ‘should be fated to constitute IPs on their own’. Other authors have linked phrasing to polarity. According to Gussenhoven (1984: 118–20; 2004: 292), reversed polarity tags consist of an accented auxiliary and a pronominal subject and have to be phrased separately, while constant polarity tags incorporate obligatorily: they are ‘always included in the tail of the nuclear tone in the host sentence’ (Gussenhoven 1984: 120).

There is also disagreement in the literature regarding the intonational realization of QTs. A number of researchers have noted that pitch accent choice on the QT is related to polarity and the interpretation of the utterance (e.g. Crystal 1969: 273; Sadock 1974: 126–34; Rando 1980: 245–7; Gussenhoven 1984: 118f.; Bing 1985: 36; Quirk *et al.* 1985: 811–12; Huddleston & Pullum *et al.* 2002: 894–5). A reversed polarity tag has been argued to express an expectation (neutral, negative or positive) about the assumption expressed in the host. Depending on its interpretation, it may have either rising or falling intonation: The rising contour has been argued to be interrogative and at the same time to have a hedging function; it is used to express uncertainty or doubt on the part of the speaker about the truth of the proposition expressed in the host, and to ask for verification. The falling contour, on the other hand, is used to express a speaker's certainty about the truth of the host's proposition and is confirmation-seeking; it invites confirmation or acknowledgement and is closer to a statement than a question. While falling and rising tones have been observed for reversed polarity tags 'independently of the nuclear contour on the preceding [domain]', it has been argued that they cannot be produced with a fall-rise (H^*LH ; see Gussenhoven 2004: 292). According to Gussenhoven (1984: 118–20), constant polarity tags are statements on the part of the speaker: 'the speaker draws a conclusion on the basis of (non-) linguistic information just received' (Gussenhoven 1984: 120). The positive constant polarity tag has typically a (slightly) rising tone and may carry emotive meaning (Quirk *et al.* 1985: 812; Huddleston & Pullum *et al.* 2002: 894–5), and 'it sometimes has no nucleus, but is part of the preceding tone unit' (Quirk *et al.* 1985: 812). Bald (1980) finds a principal difference in prosodic realization between reversed polarity tags and constant polarity tags such that reversed polarity tags are predominantly falling and have a strong tendency to be phrased separately, while constant polarity tags are predominantly rising and integrated in the preceding prosodic phrase.

Unlike the approaches outlined above, which make a predominant distinction between falling and rising pitch movements, often without paying particular attention as to whether the QT is prominent or not, Ladd (1981) considers 'the distinction to be primarily one of accent and phrasing . . . rather than a question of rising or falling pitch' (Ladd 1981: 167). His basic distinction is between nuclear and post-nuclear tags. While nuclear tags have a main accent (nucleus) of their own and are generally preceded by a phrasal prosodic boundary, post-nuclear tags do not have a nucleus or preceding boundary; instead, the pitch contour continues from the preceding nucleus in the same domain, which is associated with a syllable in the host. This intonational distinction is then related to general syntactic and pragmatic distinctions (see Sadock 1974: 133; Ladd 1981: 168). In pragmatics, they argue, the nuclear tag states or asserts the speaker's assumption and signals a hedge to the proposition expressed by the host, while the post-nuclear tag is closer to true interrogatives and expresses real doubt or uncertainty. In conclusion (Ladd 1981: 168–9), the nuclear tag is separate in its meaning and intonation: it expresses a separate speech act and has its own intonation domain and contour; the post-nuclear tag, on the other hand, is integrated in meaning and intonation. If both Ladd (1981) and Gussenhoven (2004) are right, then constant polarity tags can

never express a separate speech act because, according to Gussenhoven (2004), they are always integrated. Reversed polarity tags, on the other hand, always express separate speech acts because they have to be phrased separately.

The focus of the present article will be on the prosodic phrasing and intonational realization of QTs in English and on factors affecting their prosody.³ Based on previous literature, we assume that QTs are not phrased separately across the board (see also previous findings on prosodic integration of other types of short, light parentheticals; e.g. Dehé 2009 and references therein), as would follow from a one-to-one mapping of syntactic parenthetical status to prosodic constituency (e.g. Knowles 1980; Selkirk 1984; Potts 2002). Instead, the prosodic realization of QTs is influenced by polarity, syntactic position, position in a speaker's turn and verb type, along with pragmatic factors (not addressed here). We tested the hypotheses in (5) and (6), which have not yet been systematically studied on the basis of a large data set taken from actual spoken language.

(5) Hypotheses: prosodic phrasing

- (a) Prosodic phrasing is determined by polarity such that reversed polarity tags are obligatorily phrased separately, while constant polarity tags must incorporate (see Gussenhoven 1984, 2004).
- (b) Prosodic separation of the QT is more frequent when the QT is located at a syntactic boundary than when it occurs phrase (XP)-medially. This follows from standard assumptions about the syntax-prosody interface (e.g. Selkirk 1984, 1986, 1995).
- (c) Prosodic phrasing is affected by verb type (Bolinger 1989), such that QTs containing modals are typically integrated.

(6) Hypotheses: intonational realization

- (a) QTs never precede the nuclear accent; they either have a nuclear accent, or they are (part of) post-nuclear material within an IP (Ladd 1981).
- (b) QTs may not have falling-rising contours, but only falling or rising pitch contours (Gussenhoven 2004).
- (c) In nuclear QTs, the accent is never realized on the pronoun (Knowles 1980; Quirk *et al.* 1985).
- (d) The intonational realization of QTs is related to polarity. Constant polarity tags have rising contours (Bald 1980; Gussenhoven 1984). Reversed polarity tags may either be rising or falling (Gussenhoven 1984) or they are predominantly falling (Bald 1980).
- (e) We expect more falling contours within a speaker's turn, expressing certainty and not asking for an immediate reply. No firm predictions are possible for QTs at the end of a turn as these are dependent on a larger range of pragmatic factors.

³ Most of the work on QTs referred to above relates to intonation domains in the British system of intonation analysis, i.e. without assuming a hierarchy of two phrasal levels above the prosodic word, or it relates to IPs in the autosegmental-metrical (AM) system. The intonation domain and the IP have in common that they are the domain of a complete tonal contour/nuclear accent (e.g. a nuclear fall-rise in the British system, corresponding to a nuclear H* L-H% in the AM-system). Note that Beckman & Pierrehumbert (1986: 295) take the boundary between host clause and QT as one piece of evidence for the existence of the intermediate phrase (ip) in English. In our analysis, 'separate' (= non-integrated) QTs all have their own complete tonal contour/nuclear accent and are set off from previous material by a boundary tone, often accompanied by a pitch discontinuity and/or pause. They are thus intonation domains/IPs in their own right (see also section 2.2 below).

In addition, we tested whether the intonational realization of the host clause affects the intonational realization of the QT.

In order to address these hypotheses, several hundred QTs were extracted from the ICE-GB (see section 2), and were subsequently prosodically analysed. Section 2 of this article describes the data and then details the linguistic and prosodic analysis. In section 3, we report the results of the prosodic analysis and the factors that influence the prosodic phrasing and intonational realization of QTs in English. In section 4, the results will be discussed in light of the hypotheses before we conclude the paper and raise issues for future research (section 5).

2 Description of the data set

The data for this study were retrieved from the spoken part of the British Component of the *International Corpus of English* (ICE-GB; Releases 1 and 2, ICE-CUP versions 3.0 and 3.1; see Nelson *et al.* 2002 and the ICE-GB website), using an automatic search followed by manual sorting. The spoken part of this corpus contains 637,682 words from various text categories, ranging from casual, informal speech such as direct casual conversations and private telephone calls, to more formal speech, including public dialogues and discussions, unscripted and scripted monologues and broadcast news, all recorded in the 1990s. The ICE-GB corpus and accompanying materials provide detailed information on the sources of the corpus data (see Nelson *et al.* 2002: appendix 2, in particular). The spoken part of the ICE-GB corpus contains accompanying sound files, which were used for prosodic analysis. The corpus is fully parsed syntactically, but not prosodically, except that pauses (i.e. silent intervals of all kinds) are indicated in the corpus transcriptions.⁴

2.1 Lexical and distributional analysis of QTs

The corpus was automatically searched for the function TAGQ (tag question; Nelson *et al.* 2002: 55), yielding 722 QTs in 718 utterances. Of these, 338 utterances (339 QTs) had to be discarded in the manual sorting process. They either had no available sound file or the quality of the sound files was not good enough for prosodic analysis (overlapping speech, laughter, noise or music in the vicinity of the QT, too noisy or too quiet recordings, or items that were trimmed at locations that are important for the current analysis). Furthermore, QTs were disregarded if the pronominal subject or the auxiliary in the tag did not correspond to subject or verb in the host, respectively (e.g. (7) and (8)), if the tags were of a different syntactic form (e.g. the interrogative in

⁴ Silent intervals are indicated by one or more commas between angled brackets (<,>, <.,>, etc.); the number of commas indicates the perceived (but not measured) duration of the silence. Since silent intervals do not automatically correspond to boundaries of phrasal prosodic constituents, these pause indicators were removed and do not therefore occur in any of the corpus examples given throughout the article, even if present in the original corpus transcriptions. After removal of the pause indicators, all data appear without punctuation.

Table 1. *Distribution numbers of QTs according to auxiliary verb in the tag and polarity type (reversed polarity tags: pos-neg, neg-pos; constant polarity tags: pos-pos, neg-neg)*

	<i>be</i>	<i>do</i>	<i>have</i>	Modal	Overall
pos-neg	149	68	26	28	271
neg-pos	26	25	5	12	68
pos-pos	26	11	0	5	43
neg-neg	1	1	0	0	2
Overall	202	105	31	45	383

(9)) or if they had been misannotated as tags. The remaining 380 utterances containing 383 QTs entered the analysis. Table 1 shows a cross-tabulation of QTs according to auxiliary verb and polarity. In the data set analysed here, reversed polarity tags with negation in the tag (pos-neg) are most frequent across all verbs, followed by reversed polarity tags with negation in the host. Furthermore, *be* is the verb most frequently found in QTs, followed by the verb *do*, while QTs with *have* and modal verbs are less frequent. This distribution is consistent with Tottie & Hoffmann's (2006: 290–1) findings for British and American English. Examples from the present data set are given in (10); (10c) is one of the utterances containing two QTs.

- (7) Example QT excluded from analysis because of lack of correspondence between the subjects in the QT and the host:
 Brandon's going to lend me loads of money *aren't you* cos my cash and money are stolen (ICE-GB: s1a-040 #368)
- (8) Example QT excluded from analysis because of lack of correspondence between the verbs in the QT and the host:
 Oh that'll be handy *wouldn't it* (ICE-GB:s1a-019 #241)
- (9) Example QT excluded from the analysis because of syntactic form:
 And you suffer from mild asthma *is that right* (ICE-GB: s1a-051 #106)
- (10) Example corpus items containing QTs (from the data set):
- (a) 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 uh it pays off (ICE-GB: s2a-031 #080)
- (b) And uh we've learnt *haven't we* that books are all right but there is nothing like uh the help you get from society members as the team have said and doing it yourself (ICE-GB: s1b-025 #166)
- (c) And you were satisfied with those *weren't you* because you already had the earlier accounts *didn't you* (ICE-GB: s1b-065 #127)

Furthermore, the QTs were coded according to their syntactic position and their position in a speaker's turn. A QT was coded as utterance-final (see (11a)) if it occurred at the end of a syntactic sentence and simultaneously at the end of a turn, as sentence-final (see (11b)) if it was positioned at the end of a syntactic sentence (followed by a full stop

in the corpus transcriptions) but the same speaker continued in the discourse. Further positions were phrase (XP-) final (see (11c)) and XP-medial (see (11d)).

(11) Position

(a) Utterance (U-) final

Speaker A: but you've never touched that form *have you*

Speaker B: No because I never liked it

I never liked reading it I mean (ICE-GB: s1b-048 #125–127)

(b) Sentence (S-) final

Well it's a very boring sort of cheese *isn't it*

I mean it's all right thinly sliced with a with an apple I suppose (ICE-GB: s1a-061 #354–355)

(c) Phrase (XP-) final

He suffered great mental distress *didn't he* after the war (ICE-GB: s1b-032 #164)

(d) XP-medial

(d1) but isn't that at the root of it because 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 the possibility *isn't there* of lucrative television work (ICE-GB: s1b-050 #072)

(d2) You don't *do you* want us to think of faith as a synonym for tradition (ICE-GB: s1b-028 #103)

2.2 Prosodic analysis of QTs

The 383 QTs were analysed auditorily, and, if possible, instrumentally. Items whose sound files were not of good enough quality for instrumental analysis were excluded only if the auditory analysis was inconclusive. In the auditory analysis, the relevant locations and types of pitch accents and intonation phrase (IP) boundaries and boundary tones were identified. The annotations include edge tones of intermediate phrases in combination with IP boundary tones, but the analysis of prosodic phrasing was restricted to the IP level and did not include the intermediate phrase level, which is harder to annotate reliably (Syrdal & McGory 2000). Therefore, QTs labelled separate here have a clearly marked IP boundary (see (12) below for detailed criteria).

For the purpose of the instrumental analysis, the sound files were edited into individual files containing the QT, its host, and as much following material as necessary to determine the right-hand IP boundary. The instrumental analysis was done in Praat (Boersma & Weenink 2012) following ToBI (Silverman *et al.* 1992; Beckman & Ayers Elam 1993). We identified the tonal contours associated with the QT and the material in its immediate environment (the host clause). More specifically, we coded the nuclear tunes according to whether they were falling (e.g., H* L-L%), falling-rising (e.g. H* L-H%) or rising (e.g. L* L-H%, L*+H H-H%). If boundaries could not be identified immediately before and/or after the QT, the positions of the closest boundaries to the right and left of the QT were marked and the occurring pattern was described. The criteria for IP boundaries followed those discussed in the literature (see (12)).

- (12) Criteria for the identification of an IP/intonation domain
- (a) Domain-internal criteria
 - (a1) complete tonal contour (CTC)
 - (a2) domain across which declination applies
 - (a3) creaky voice
 - (b) Criteria at a potential boundary
 - (b1) presence and nature of pauses (structure-related vs hesitational)
 - (b2) pitch on unaccented syllables following a nuclear tone
 - (b3) domain final lengthening
 - (b4) presence or absence of segmental processes

The domain-internal criteria in (12a) are tonal and laryngeal features which apply within the target domain. The IP/intonation domain is defined intonationally in terms of a complete tonal contour (CTC), i.e. a nuclear tone according to the British tradition (e.g. fall, fall–rise, rise, rise–fall), or a combination of nuclear pitch accent and boundary tone in the autosegmental-metrical model (i.e. ToBI; e.g., H* L-L%, H* L-H%, L*+H H-H%, L*+H L-L%). The IP/intonation domain is also the domain across which declination applies, such that later pitch peaks within an IP are realized with lower pitch compared to earlier ones, and pitch reset indicates a new domain (see Ladd 1986, 1996/2008). However, pitch peaks in later domains may not be reset to the same maximum level as pitch peaks in earlier domains (Ladd 1986; but see Truckenbrodt 2002). Furthermore, the use of creaky voice has been considered a boundary marker at the end of an IP (e.g. Lehiste 1973; Kreiman 1982).

The criteria listed in (12b) apply at boundaries between IPs/intonation domains. Although domains may be separated by pauses (e.g. Nespor & Vogel 1986: 188), pauses are not obligatory and their occurrence, length and position depend on a number of factors (e.g. Krivokapić 2007). In the present context, all pauses were identified and carefully studied as to their nature. In particular, structure-related breath pauses, but not pauses due to performance factors such as hesitational stretches, were interpreted as boundary markers. Furthermore, the pitch on unaccented syllables following a nuclear accent was studied. Pitch discontinuities, a step-up in pitch on an unaccented syllable after a nuclear fall, and a step-down in pitch on an unaccented syllable following a rise, were seen as evidence for a boundary (Cruttenden 1997: 34). In addition, final lengthening, in particular, lengthening of the very last syllable of a domain, lengthening of the last stressed syllable, and lengthening of the entire last word in a phrase, was interpreted as a signal for a boundary (Lehiste 1973; Vaissière 1983, Gussenhoven & Rietveld 1992; Wightman *et al.* 1992; Ferreira 1993; Turk & Shattuck-Hufnagel 2007, among many others). However, final lengthening was only an impressionistic factor that could not be systematically analysed due to the (uncontrolled) nature of the corpus data. The same is true for the presence or absence of segmental processes (see also Knowles 1991: 155–8).

All items were analysed by the first author. Cases that caused uncertainty were given to at least one other experienced linguist trained in intonational analysis. Thirteen items for which a final analysis could not be agreed upon or which were indistinct were not analysed further. This left 370 QTs for analysis.

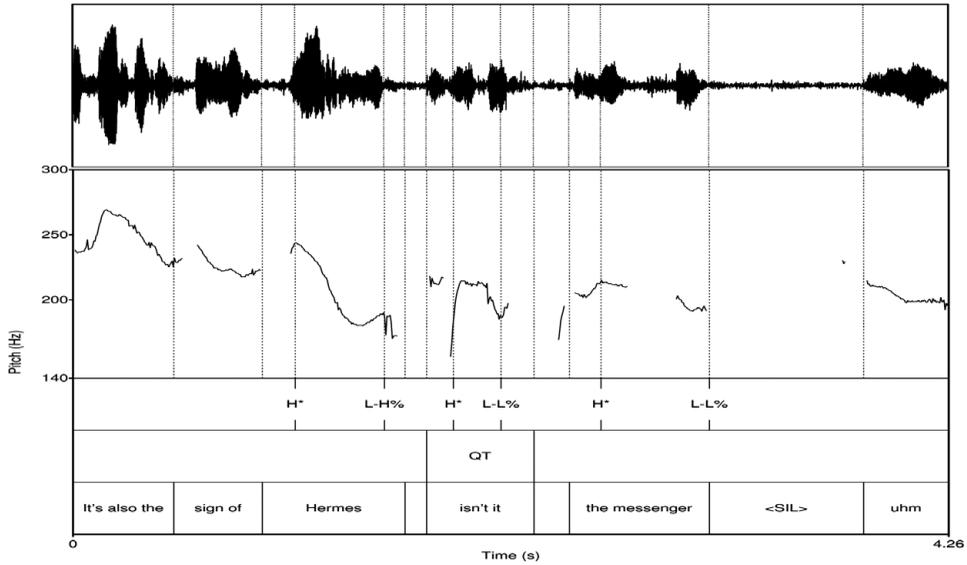


Figure 1. Phrasing pattern (13a1): QT separate (IP[...]IP IP[QT]IP IP[...]IP) ...; example (14); QT associated with falling accent

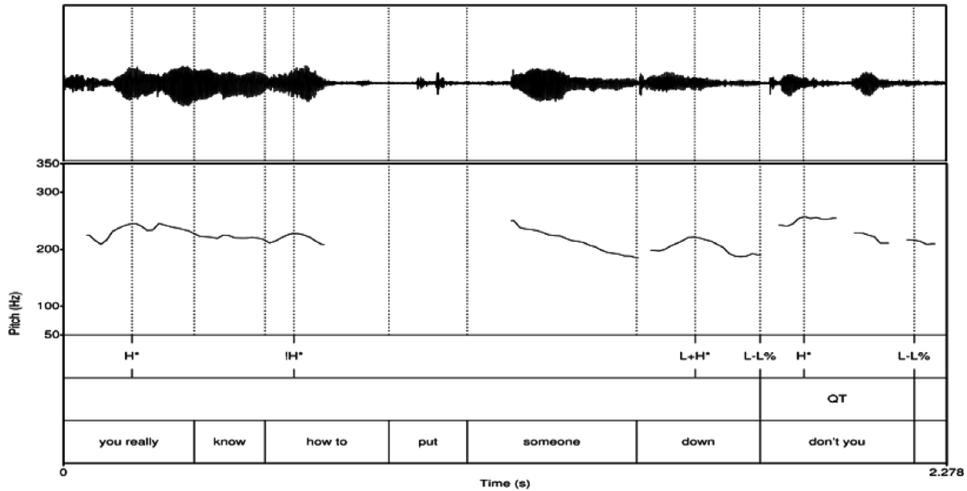


Figure 2. Phrasing pattern (13a1): QT separate (IP[...]IP IP[QT]IP); example (15), QT associated with falling accent

- (15) QT final, separate, falling accent
 - (a) God you really know how to put someone down *don't you* (ICE-GB: s1a-038 #252)
 - (b) IP1[you really know how to put someone down]IP1 IP2[don't you]IP2
 H* !H* L+H* L-L% H* L-L%

In (16)/figure 3, IP1 and IP2 have CTCs: a falling accent (H+!H* L-L%) in IP1 preceding the QT and a rising accent (L*+H H-H%) in IP2 (the QT *hasn't it*). The fall

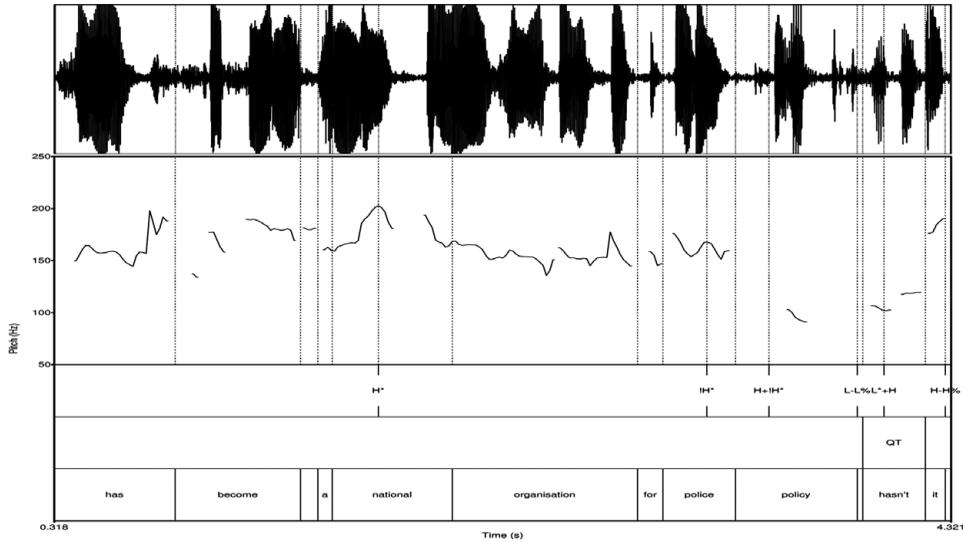


Figure 3. Phrasing pattern (13a1): QT separate ($IP[\dots]_{IP} IP[QT]_{IP}$); example (16); QT associated with rising accent

to L-L% in IP1 reaches the lowest point in the speaker's pitch range in this utterance; it is accompanied by (impressionistic) domain-final lengthening and creaky voice and there is a discontinuity before the beginning of the tag. The prominence and rising accent of the QT is presumably due to its information-seeking nature. It is truly interrogative and the interlocutor replies immediately (see (16c)).

(16) QT final, separate, rising accent

- (a) Speaker A: uh ACPO as it's called uh accountable to nobody has become a national organisation for police policy *hasn't it* (ICE-GB: s1b-033 #69)
- (b) IP_1 [has become a national organization for po lice policy] $_{IP_1}$ IP_2 [hasn't it] $_{IP_2}$
H* H* H+!H*L-L% L*+H H-H%
- (c) Speaker C: I wouldn't accept that ACPO is not specifically accountable to nobody as to use your phrase [. . .] (ICE-GB: s1b-033 #70–71)

In (17)/figure 4, the QT has a falling–rising pitch accent and its domain (IP2) is terminated by L-H%; it is preceded by a short pause and followed by a longer pause (> 500 ms). IP1 has a falling nuclear accent associated with *model* and terminates in L-L%. The F0 peak on the tag is slightly lower than the final peak in IP1.

(17) QT non-final, separate, falling–rising accent

- (a) I mean actually you've therefore taken a model *haven't you* which you produced <unclear-words> (ICE-GB: s1a-064 #047)
- (b) IP_1 [I mean actually you've therefore taken a model] $_{IP_1}$ IP_2 [haven't you] $_{IP_2}$
H* !H* L+H* L+H* L-L% H* L-H%

In accordance with hypothesis (6c), we find that the accent is never associated with the pronoun. In 275 of the 278 QTs, the accent is realized on the auxiliary; the

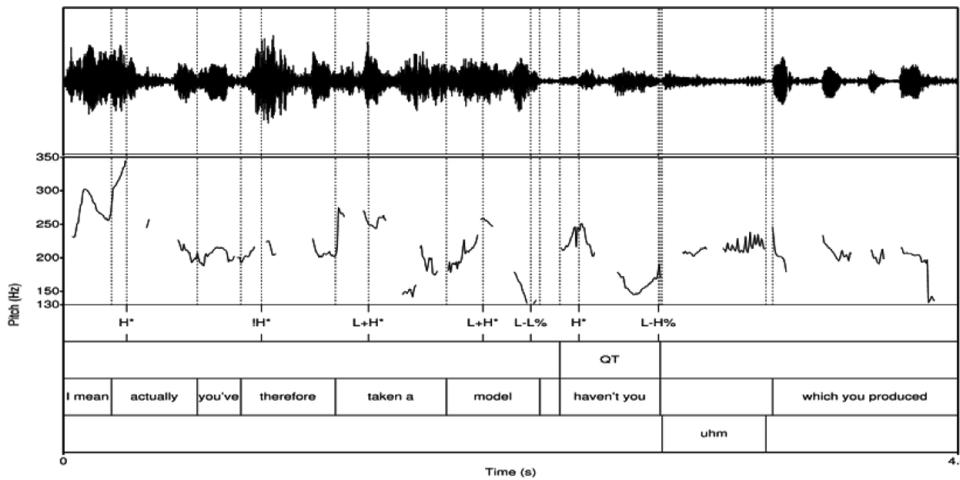


Figure 4. Phrasing pattern (13a1): QT separate ($[IP[\dots]IP[QT]IP[\dots]IP]$); example (17), QT associated with falling–rising accent

remaining three QTs all have a negative tag (one neg-neg, two pos-neg; see (18)) and the accent is associated with the negation *not*. In (18a) the tag is preceded by L-L% and a silent pause. It has a rising accent associated with *not*. Semantically, the tag is interrogative and a reply is given by the interlocutor (see the excerpt of the dialogue in (19)). Similarly, (18b) has a rising accent on *not* asking for confirmation, and the interlocutor immediately replies *Yes* before the speaker continues. (18c) has a falling accent on *not* and the same speaker continues, not asking for a reply.⁶

(18) QTs phrased separately, main prominence on *not*

- (a) 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* (ICE-GB: s1b-064 #113)
- (b) Uh didn't Mr Hook say to you that at this point that he was interested in further expansion of the uh Ferndale business *did he not* (ICE-GB: s1b-064 #118)
- (c) Now in practice there was censorship *was there not* by which it was I I a suggested and b very readily and compliantly agreed that nothing that might uh disturb the patriotic mood over the conflicts should be allowed to appear on patriotic <unclear-word> newspapers (ICE-GB: s1b-031 #055)

(19) Dialogue context of example (18a)

Speaker A: 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* (ICE-GB: s1b-064 #113)

Speaker B: Well Hook said one thing and uh the other guy said something else. They both were at almost at conflict with each other (ICE-GB: s1b-064 #114–115)

Speaker A: Well let me concentrate on what Mr Hook said (ICE-GB: s1b-064 #116)

⁶ It is possible, of course, that there is a non-verbal reply on the part of the interlocutors.

Two QTs with *do* (one pos-neg, one neg-pos) were phrased in a Compound Prosodic Domain (CPD; Ladd 1986, 1996/2008) pattern, a structure analogous to a compound in morphology (see (13a2)).⁷ An example is given in (20). There is no reason to assume an IP boundary after *don't*, preceding the tag. IP1 is the domain of declination and its contour continues after the insertion of the tag (and the L+!H* on *want* is therefore downstepped with regard to the H* on *don't*). The QT *do you* is prominent and has its own CTC: a falling–rising pitch accent associated with *do*. Its pitch peak is lower than in the surrounding IP, but it is perceptibly not subordinated to the main accent in IP1. It is conceivable that the phrasing pattern is related to the syntactic position of the tag: it occurs within the verbal complex between auxiliary and main verb, which is not a position at which a strong prosodic boundary would be expected.

(20) QT phrased in Compound Prosodic Domain

(a) You don't *do you* want us to think of faith as a synonym for tradition
(ICE-GB: s1b-028 #103)

(b) IP1[you don't IP2[do you]IP2 want us to think of faith]IP1
H* L+H*L-H% L+!H* !H* L-L%

3.1.2 *Pattern (13b): QT followed (but not preceded) by material from the host in its IP (... IP[...]IP IP[QT ...]IP (IP[...]IP) ...)*

Six QTs were prominent but were joined in their domain by material from the host, corresponding to phrasing pattern (13b). Of these, two QTs were followed in their domain by unstressed vocatives (e.g. (21)), two by adverbs (one by *here*, see (22a), and one by *really*, see (22b)) and two by hesitant stretches or discourse markers such as *I mean* (e.g. (23)). All six QTs have a falling accent with the nuclear peak associated with the auxiliary and the falling/low contour being continued on the material following the QT in the same domain (see (21b) and (23b)). In (23), *escape* is lengthened and IP1 is terminated by L-L%. *I mean* continues the fall/low pitch on the QT and is followed by a step-up in pitch on *that's*. After the second *that's* the interlocutor steps in (*Oh yes. Yeah I mean you're escaping into the book in uh you know everything around you;* ICE-GB: s1a-016 #325–327).

(21) QT phrased according to (13b): QT followed by vocative in its IP

(a) and then I say well next week is a full week's block teaching *isn't it Paul* (ICE-GB: s1a-082 #42)

(b) IP1[... block teaching]IP1 IP2[isn't it Paul]IP2 IP3[...]IP3
H* L* L-H% H* L-L%

⁷ The Compound Prosodic Domain (CPD) is defined as a 'prosodic domain of a given type X whose immediate constituents are themselves of type X' (Ladd 1996: 244). In the present context, a CPD is a prosodic domain of type IP, whose immediate constituent – the QT – is itself of type IP. The idea of the CPD was based, among other evidence, on the observation that a parenthetical inserted into a host may interrupt the declination of the host, which is then resumed after the completion of the parenthetical (Cooper & Sorensen's 1981 findings; see Ladd 1986, 1996 for discussion). The host spans one domain, which is interrupted by another domain (the parenthetical; the QT in the present context), yielding a recursive prosodic structure: the CPD (IP1[host IP2[QT]IP2 host]IP1). Since in a CPD the QT has its own IP we count it as separate.

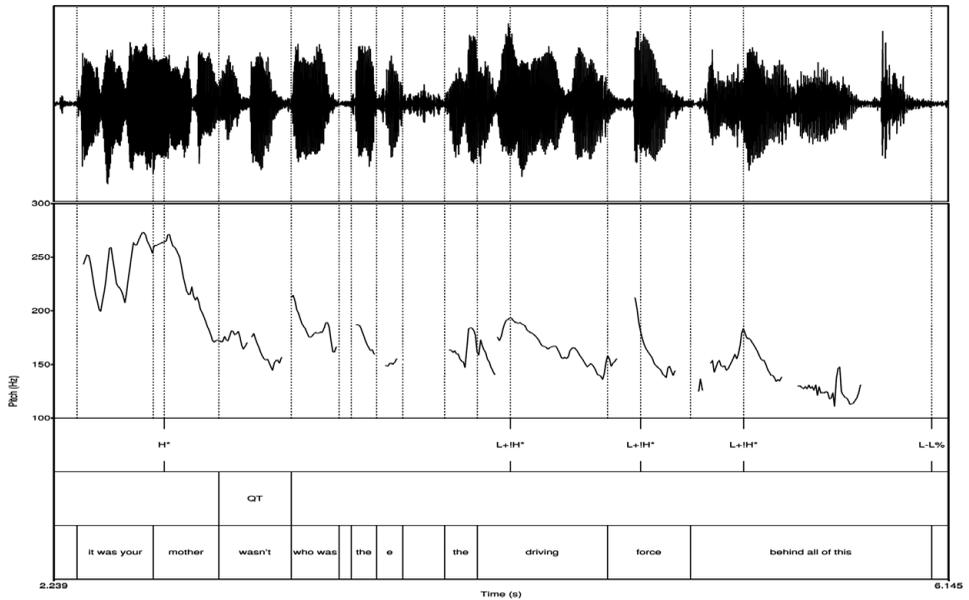


Figure 6. Phrasing pattern (13c2); QT integrated, i.e. phrased together with preceding and following host material ($_{IP}[\dots QT \dots]_{IP}$); example (26), QT preceding the nuclear syllable as part of the transition between two prenuclear accents

A QT with a falling contour is shown in (25), the nuclear accent is associated with *Xepe*, the fall is continued on the particle verb *turned up* and the QT (and its IP) is terminated by L-L%.

- (24) QT realized as unstressed, post-nuclear material, rising
 (a) Oh really, can't fit me in as well *can you* (ICE-GB: s1a-039 #360–361)
 (b) $_{IP1}[\text{can't fit me in as well can you}]_{IP1}$
 H* !H* L-H%
- (25) QT realized as unstressed, post-nuclear material, falling/low
 (a) Oh *Xepe* turned up *did he* (ICE-GB: s1a-005 #139)
 (b) $_{IP1}[\text{Oh Xepe turned up did he}]_{IP1}$
 L+H* L-L%

Finally, five QTs were phrased according to (13c2), i.e. they were preceded and followed by material from the host in the same IP. One of the five was unstressed and preceded the nuclear accent in its domain (see (26)/figure 6); this is the only exception to the claim put forward in hypothesis (6a) (Ladd 1981) that QTs never precede the nuclear accent.

- (26) QT realized as unstressed material preceding the nuclear accent
 (a) So on going back to your to your childhood it was your mother *wasn't it* who was the the driving force behind all of this behind this sort of intellectual rigour (ICE-GB: s1b-046 #057)

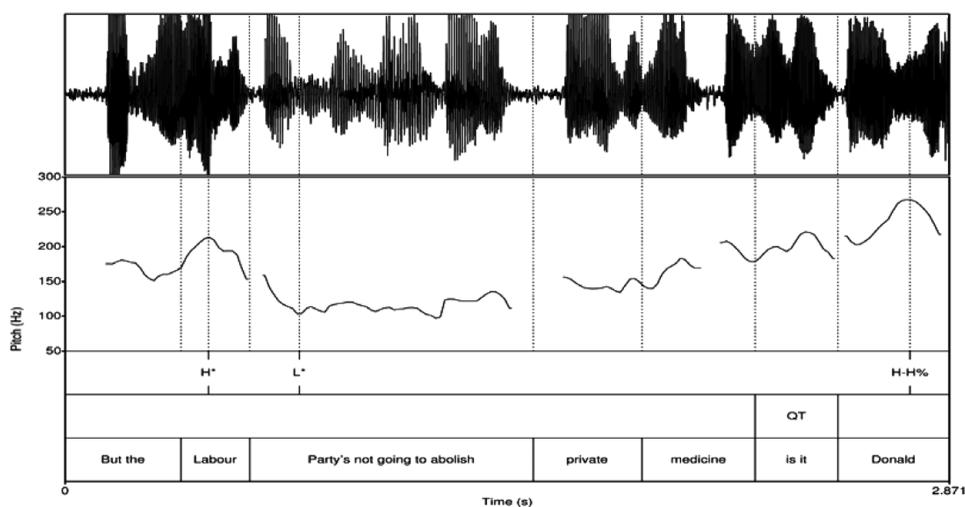


Figure 7. Phrasing pattern (13c2); QT integrated, i.e. phrased together with preceding and following host material ($IP[\dots QT \dots]IP$); example (27), QT following the nuclear syllable as part of post-nuclear material

- (b) IP_1 [So on going back to your to your childhood] IP_2 [it was your mother wasn't it who was the the driving force behind all of this] IP_3 [behind this sort of intellectual rigour] IP_3
- (c) [it was your mother wasn't it who was the the driving force be hind all of this]
 H^* $L+!H^*$ $L+!H^*$ $L+!H^*$ $L-L\%$

The utterance in (26) is divided into three rather long IPs (see (26b), all of which have a CTC, are the domain of declination and are terminated by L-L%. The tonal analysis of the IP containing the QT (IP_2) is given in (26c)/figure 6. The QT *wasn't it* is unstressed and reduced to /wɒz²nt/, i.e. the pronoun disappears.⁸ It is realized in the transition between two prenuclear accents associated with *mother* and *driving*, respectively. The nuclear accent of this IP is associated with *behind*.

The other four instances in this phrasing condition were unstressed and followed the nuclear syllable; they were either rising (three items; e.g. (27)/figure 7) or falling (one item). Similarly to (21), the QT in (27) is followed by a vocative, but unlike the QT in (21), the QT in (27) is not prominent. In a long IP, which spans the whole utterance, the QT is part of the rising post-nuclear material in a falling–rising pattern; the nuclear pitch accent appears early in the host. Auditory and visual inspection of the data gives no reason to assume an IP-boundary anywhere else.

⁸ Wichmann (2007) notes that QTs might be most likely to be reduced if they 'have no informational or discourse-organising function' (Wichmann 2007: 356), but have an attitudinal function instead. Attitudinal QTs, she assumes, are most probable utterance-medially (Wichmann 2007: 350). The QT in (26) is compatible with Wichmann's analysis; however, a thorough analysis of all QTs in the corpus relating their prosody to interpretation and function in discourse has yet to be done.

Table 2. *Cross-tabulation of QTs according to phrasing and intonation contour*

Phrasing	Falling contour on QT	Rising contour on QT
Separate (with IP phrase break)	243	20
Integrated (without IP phrase break)	5	61

(27) QT realized as unstressed post-nuclear material

(a) But the Labour Party's not going to abolish private medicine *is it* Donald
(ICE-GB: s1b-039 #88)

(b) _{IP1}[But the Labour Party's not going to abolish private medicine *is it* Donald]_{IP1}
H* L* H-H%

3.2 Factors affecting the prosodic phrasing and intonational realization of QTs

Following the descriptive analysis, we analysed the data statistically to test which factors have an effect on the prosodic phrasing and intonational realization of QTs in English. The data were coded according to the factors polarity (pos-neg, neg-pos, pos-pos), position (utterance-final (turn-final), sentence-final (turn-medial), phrase(XP-final, XP-medial) and verb type (*be, do, have, modal*).

Of the 370 QTs that were conclusively analysed and included in the descriptive analysis reported in section 3.1 above, 41 had to be discarded from the statistical analysis for the following reasons: the only two cases with constant negative polarity tags (neg-neg) and six cases with a nuclear accent on the QT (phrased as in (13b)/section 3.1.2) were excluded because of too few data points. The same was true for 24 cases with infrequent intonation contours on the tag (8 with falling–rising, 15 with low and one unaccented). Finally, nine cases in which the syntactic position was unclear, mostly due to hesitant stretches or self-corrections on the part of the speaker, were not considered.

The remaining 329 cases entered the statistical analysis. The dependent binary variables were prosodic phrasing (integrated; see (13c)/section 3.1.3 and non-integrated (= separate); see (13a)/section 3.1.1), and intonational contour on the QT. The QT could be realized falling/low (in separate QTs e.g., with H* L-L%, in integrated QTs with a low or falling pitch, leading to a low boundary tone) or rising (in separate QTs e.g., with L*(+H) H-H% or L* L-H%, in integrated QTs with a rising pitch leading to a high boundary tone).

The prosodic annotation suggested a correlation between prosodic phrasing and intonational realization of the QT. As can be seen in table 2, separate QTs are generally realized with a falling contour, and integrated QTs with a rising one (more than 90% of the cases each). This correlation is corroborated by a non-parametric Kendall's tau test (R version 2.12.2). Not surprisingly, the results show a highly significant correlation between phrasing and intonation contour (tau = 0.79, z = 14.3, p < 0.0001).

3.2.1 Phrasing

We next investigated the effects of polarity, position, and verb type, testing the hypotheses in (5). Intonation contour was not included because of its strong correlation

Table 3. *Cross-tabulation of QTs according to phrasing and polarity*

Phrasing	neg-pos	pos-neg	pos-pos
Separate	46	203	14
Integrated	18	28	20
Ratio separate : integrated	2.6	7.3	0.7

Table 4. *Cross-tabulation of QTs according to phrasing and position*

Phrasing	U-final	S-final	XP-final	XP-medial
Separate (with phrase break)	129	102	24	8
Integrated (no phrase break)	38	20	2	6
Ratio separate : integrated	3.4	5.1	12	1.3

with phrasing. The factors influencing the intonation contour of the QT will be investigated in section 3.2.2, according to and in the order of hypotheses (6). We calculated a binomial logistic regression model with phrasing as the dependent variable and the above-mentioned factors as fixed factors⁹ (function *glm* with a binary linking function, see Baayen 2008: chapter 6.3.1). To test the validity of the model, data points with residuals greater than 2.5 standard deviations were removed and the model was refitted (removing not more than seven QTs, i.e. less than 2 per cent of the data). The statistical analysis reported here is based on this final model.

Regarding polarity, QTs with reversed polarity have a higher proportion of separate compared to integrated QTs, while QTs with constant (positive) polarity are more often integrated than separate (neg-pos vs pos-pos: $\beta = 1.0$, $z = 2.0$, $p < 0.05$; pos-neg vs pos-pos: $\beta = 2.3$, $z = 5.2$, $p < 0.0001$). Furthermore, within the group of reversed polarity tags, pos-neg QTs were significantly more frequently phrased separately than QTs with neg-pos ($\beta = 1.3$, $z = 3.4$, $p < 0.0005$). Distribution numbers and ratios are given in table 3. They confirm that hypothesis (5a), claiming obligatory separation of reversed polarity tags, is too strong in its current form, but that there is a clear tendency in this direction. What is more, the reversed polarity tags do not behave uniformly in this respect.

Position also had a significant effect on the phrasing of QTs (see table 4). There were more separate QTs in utterance-final, sentence-final and XP-final position than in XP-medial position (U-final vs XP-medial: $\beta = 1.3$, $z = 2.0$, $p < 0.05$; S-final vs XP-medial: $\beta = 1.8$, $z = 2.6$, $p < 0.01$; XP-final vs XP-medial: $\beta = 2.0$, $z = 2.1$, $p < 0.05$). There was no difference in prosodic phrasing for phrase-final QTs (utterance-final, sentence-final and XP-final QTs, all p-values for all pairwise comparisons > 0.15). These results clearly confirm the syntax-prosody relation as put forward in hypothesis (5b).

⁹ Note that the distribution of the data points across factors did not allow us to test for interactions between factors.

Table 5. *Cross-tabulation of QTs according to phrasing and verb type*

Phrasing	be	do	have	modal
Separate (with phrase break)	124	84	22	33
Integrated (no phrase break)	46	8	3	9
Ratio separate: integrated	2.7	10.5	7.3	3.7

Finally, we found a significant effect of verb type on the number of separate QTs (see table 5): QTs containing the verb *do* were significantly more often separate than QTs with the verb *be* (beta = 1.5, $z = 3.5$, $p < 0.001$) and QTs with modal verbs (beta = 1.2, $z = 2.0$, $p < 0.05$); there was no difference between QTs containing *do* and *have* ($p > 0.3$). No other differences were significant (all p -values > 0.2). These results support Bolinger's (1989) hunch about verb type generally affecting phrasing of the QTs, but do not confirm his observation about a predominant integration of modals.

One reviewer suggests that the word prosodic structure of the QT (bisyllabic, e.g. *don't you*, or trisyllabic, e.g. *haven't they*) might influence phrasing more than verb type. In order to test this hypothesis, we fitted an alternative model in which the number of syllables in the QT (two or three) replaced verb type. Results showed no effect of syllable number of the QT ($p > 0.15$). The factor syllable number is also not significant if we add this factor to the main model ($p > 0.2$). To exclude the possibility that the effect of syllable number is masked by the factor polarity (all neg-pos and pos-pos QTs are disyllabic, while pos-neg QTs had both bisyllabic and trisyllabic forms), we reduced the data set to pos-neg cases ($N = 231$) and ran the model again (excluding the factor polarity). In this model also, syllable number did not show an effect ($p > 0.15$). Therefore, syllable number does not seem to be the primary factor guiding prosodic phrasing.

3.2.2 *Intonational realization*

Next, we calculated a binomial logistic regression to investigate the factors affecting the intonational realization of QTs, i.e. whether they are realized with falling or rising contours. The nuclear tune of the preceding domain (the host clause) had no effect on the intonational contour of the QT ($p > 0.25$) and was therefore removed from the model.¹⁰

As shown in table 6, polarity influences the intonational realization of QTs. Falling contours were more frequent in reversed polarity QTs than in constant polarity QTs (pos-neg vs pos-pos: beta = 3.0, $z = 6.4$, $p < 0.0001$; neg-pos vs pos-pos: beta = 1.8, $z = 3.7$, $p < 0.0005$). Within reversed QTs falling contours were more frequent in pos-neg than in neg-pos (beta = 1.1, $z = 3.3$, $p < 0.005$). These data largely confirm all

¹⁰ In total, there were 239 instances with a falling QT (209 with a preceding falling host clause and 30 with a rising one) and 20 instances with a rising QT (17 with a falling host clause and 3 with a rising one).

Table 6. *Cross-tabulation of QTs according to intonational contour and polarity*

Contour	neg-pos	pos-neg	pos-pos
Falling	44	196	8
Rising	20	35	26
Ratio of falling: rising	2.2	5.6	0.3

Table 7. *Cross-tabulation of QTs according to intonational contour and position*

Contour	U-final	S-final	XP-final	XP-medial
Falling	118	100	23	7
Rising	49	22	3	7
Ratio of falling : rising	2.4	4.5	7.7	1

Table 8. *Cross-tabulation of QTs according to intonational contour and verb type*

Contour	<i>be</i>	<i>do</i>	<i>have</i>	Modal
Falling	118	80	20	30
Rising	52	12	5	12
Ratio of falling : rising	2.3	6.7	4	2.5

aspects of hypothesis (6d), but again, reversed polarity QTs do not behave in a uniform way.

The predicted difference between utterance-final and sentence-final QTs was significant ($\beta = 0.7$, $z = 2.1$, $p < 0.05$): QTs were produced with falling contours more often sentence-finally than utterance-finally (at the end of a turn, see table 7). In addition, QTs have falling contours more frequently if utterance-final, sentence-final or XP-final, as compared to QTs in XP-medial positions (U-final vs XP-medial: $\beta = 1.3$, $z = 2.0$, $p < 0.05$; S-final vs XP-medial: $\beta = 2.0$, $z = 3.0$, $p < 0.005$; XP-final vs XP-medial: $\beta = 1.8$, $z = 2.1$, $p < 0.05$). No further significant differences were found (all p -values > 0.4). It should be noted, however, that there are very few instances in some cells. Our data fully support hypothesis (6e) in relative terms (more rises in utterance-final QTs than in other positions) but not in absolute terms (overall, more falling than rising QTs).¹¹

¹¹ For the sake of completeness, we also tested the factor verb type (see table 8). QTs containing the verb *do* have significantly more falling contours than QTs containing the verb *be* ($\beta = 1.5$, $z = 3.5$, $p < 0.0005$) and than QTs containing modal verbs ($\beta = 1.2$, $z = 2.3$, $p < 0.05$). QTs containing *do* and *have* did not differ from each other ($p > 0.1$), neither do QTs containing *be* and modal verbs ($p > 0.5$).

4 General discussion

Our results confirm the correlation between the prosodic phrasing and intonational realization of QTs, such that rising tags integrate significantly more often than falling ones. Furthermore, the factors polarity, position and verb type all have an effect to the extent outlined in section 3.2 above. This section will discuss the hypotheses formulated in (5) and (6), in turn. We will first address the prosodic phrasing of QTs, then their intonational realization.

Prosodic separation of QTs is not obligatory, despite the assumption that they are parentheticals in the syntax.¹² More than a fifth of the QTs were annotated as prosodically integrated, mostly as unstressed material following the nuclear accent in their domain. We therefore conclude that there is no strict correspondence between being a QT and prosodic separation, as suggested, for example, in Potts' (2002) theoretical account. If QTs are parentheticals in the syntax, the present findings provide additional evidence for the assumption that parentheticals, and in particular short and light types of parentheticals, are not necessarily prosodically phrased separately (see also Dehé 2009 and references given therein for comment clauses). Another interpretation is that there is no direct syntax–prosody mapping, but that the phonological form may be adjusted according to the respective formal and discourse context. This is compatible with the assumption that there is ‘no inherent relation . . . between the prosodic category types found in phonological representations and the category types of syntactic representation’ (Selkirk 2011: 437), but that instead, along with syntactic factors, there are ‘non-syntactic influences on phonological domain structure’ (Selkirk 2011: 439). However, prosodic separation is indeed the most frequent pattern of QTs in English. This may be due to their syntactic status, but also to the fact that pragmatically, they make a contribution separate from that of the host, i.e. they form a separate speech act, or a combination of the two.

Overall, prosodic separation together with integration as post-nuclear unstressed material accounts for 94 per cent of all QTs in the present data set (see (13)). This supports Ladd's (1981) observation that (non-nuclear) QTs never precede the nuclear accent in their IP; QTs are either nuclear themselves or they are (part of) post-nuclear material (see hypothesis (6a)). In terms of phrasing this means that nuclear QTs must be directly preceded by an IP boundary (Ladd 1981) and cannot be preceded by host material in their IP. This is shown in (28): the QT with nuclear prominence (indicated by underlining in (28)) is preceded by an IP-boundary. It may optionally be followed by host material in the same IP (indicated by (...) in (28)), but may not be preceded by it (see (28b); # indicates ‘not predicted’). Ladd's observation bears up against the results of the present corpus study: in accordance with (28), the present results show that QTs with a nuclear accent stand either alone in their IP (e.g. (14)–(17)), or they are followed by unstressed material such as adverbs or vocatives (e.g. (21)–(23)), both

¹² A reviewer raised the question of whether or not QTs are parentheticals. On the basis of the prosodic analysis alone, we are not in a position to make a new contribution to the debate about the (syntactic) parenthetical status of QTs.

located at a phrasal or clausal syntactic boundary or at the end of an utterance than when positioned phrase-medially. This result is compatible with standard assumptions about the syntax–prosody interface, which predict the coincidence of prosodic and syntactic boundaries, and it includes syntactic phrasing as one factor shaping prosodic constituent structure. What is surprising, perhaps, is that XP-final position did not differ from sentence- and utterance-final positions because, according to the end-based approach and subsequent Match Theory of Selkirk (1984, 1986, 1995, 2009, 2011), prosodic boundaries at clause edges are predicted to be stronger than those at XP edges. However, it is necessary to keep in mind that ‘separate’ in the present article means phrased in an IP or, in the British tradition, in an intonation domain, according to the criteria listed in (12). We did not attempt to explicitly distinguish between separate prosodic phrases of different levels in the prosodic hierarchy (e.g. in particular, intermediate phrase vs IP). It is therefore impossible to rule out that some QTs might have to be analysed as ips (in the terminology of Beckman & Pierrehumbert 1986) in a more fine-grained analysis. How, for QTs, this distinction relates to position and whether it would result in a statistically significant difference between QTs at XP-boundaries and at clausal boundaries is a topic for future research. Furthermore, the effect of position on prosodic phrasing also shows that Ziv’s (1985: 189) strong claim that QTs are ‘restricted to sentence-final position’ is untenable. The spoken corpus investigated here includes numerous counterexamples.

Third, the factor verb type also affected phrasing (see hypothesis (5c)), in line with Bolinger’s (1989: 164) assumption that the choice of the tag may affect its prosodic realization. However, while Bolinger (1989) argues that it is common for QTs containing *will* and *can* to be integrated (and he remains relatively vague otherwise), our results allow us to be more specific. In particular, the verbs *do* and *have* seem to be special such that they phrase separately more often than *be* and modals. At this point, we have no satisfactory explanation for this finding. The frequent prosodic separation of QTs containing the verb *do* may be explained by its function as ‘dummy *do*’; possibly, the need to replace a verb from the host clause by *do* is effortful and results in more prosodic separation. This explanation cannot be extended to QTs with *have*, though. This effect is not merely a reflex of one of the other factors tested as their effect was factored out of the statistical model. An alternative interpretation, that the word prosodic structure of QTs (whether they are bisyllabic or trisyllabic) affects prosodic phrasing, was not supported by the statistical analysis.

Regarding the intonational realization of English QTs, hypothesis (6a) is borne out, again in line with Ladd’s (1981) account: QTs either have a nuclear accent (QTs phrased according to (13a, b)), or they are (part of) post-nuclear unstressed material (QTs phrased according to (13c)). Only one out of 370 QTs is not compatible with hypothesis (6a); instead it is unstressed and precedes the nuclear accent in its IP. According to hypothesis (6b), QTs may not have falling–rising intonation. Our results indeed show that falling and rising patterns are more frequent, but falling–rising nuclear QTs nevertheless occurred in 2 per cent of the cases (see (17)/figure 4 in section 3.1, for an example). It is tempting to speculate that this negligible number of falling–rising

QTs merely represents speech errors, but we cannot exclude the possibility that they are functional. In case of the QT illustrated in figure 4, for example, it is possible that the final rise functions as a continuation rise: the QT is phrase-final, but followed by a modifying non-restrictive relative clause added by the same speaker. We therefore consider Gussenhoven's (2004) claim – that QTs may not have falling–rising contours – too strong. In line with hypothesis (6c), the accent in nuclear QTs is never associated with the pronoun, confirming Knowles' (1980) and Quirk *et al.*'s (1985) observation.

We furthermore found the predicted effect of polarity (hypothesis (6d)). There were more rising contours in constant polarity QTs than in reversed polarity QTs. However, no distinction has previously been made between pos-neg and neg-pos QTs within the group of reversed polarity tags. According to our results, falling contours are more frequent in pos-neg than in neg-pos environments. Future empirical data are necessary to investigate whether the effect of polarity on intonational realization (and phrasing) is genuinely pragmatic in nature in that it primarily signals the difference between different speech acts (Reese & Asher 2008).

The intonational realization of QTs is also affected by their position (hypothesis (6e)). QTs were produced with falling contours more often sentence-finally (in turn-medial position) than utterance-finally (at the end of a turn). Generally speaking, there is a tendency for speakers to use falls at the end of utterances to signal finality, and rises to mark continuation (e.g. Cruttenden 1997; Hirschberg 2004). On the other hand, rising QTs have been argued to be interrogative and to have a hedging function; they invite responses from the interlocutor and therefore occur less frequently within a turn (sentence-finally) than at the end of turns (utterance-finally).

Our analysis also shows that the intonational realization of the host does not affect the prosodic realization of the QT. This is perhaps not surprising because the majority of QTs (approximately 75 per cent) have nuclear prominence, which has been related to independent meaning and intonation (Ladd 1981). Specifically, nuclear QTs have been argued to express separate speech acts, which have their own intonation contours, independent from that of the host. Note once again that pragmatic factors and discourse effects have not been included in the analysis but are no doubt important factors in the realization of QTs. A pragmatic analysis of the data investigated here, with subsequent statistics taking all factors into account, is desirable, but only feasible with a larger data set.

5 Conclusion and outlook

This systematic analysis of corpus data reveals that several factors play a role in the prosodic phrasing and intonational realization of QTs in English. Our data corroborate some previous claims in the literature on the effects of polarity, position in the sentence and discourse, as well as verb type. Moreover, our analyses add some new aspects on the phrasing and intonational realization of QTs in English. First, we were able to specify the effect of verb type on the phrasing of QTs, showing that QTs containing the verbs *have* and *do* are more often phrased separately (compared to *be* and modal verbs). This effect of verb type cannot be recast as an effect of the QT's prosodic structure (i.e.

number of syllables). Based on the factors we tested, we have at present no conclusive explanation for this finding. It is conceivable that pragmatic factors are at stake here. We leave this issue to future research.

A second novel contribution of the current corpus study concerns the effect of polarity. While it is true that reversed polarity tags (neg-pos and pos-neg) and constant polarity tags (pos-pos) differ in both phrasing and intonational realization, our data further show that reversed polarity tags do not behave uniformly in these respects. Specifically, pos-neg QTs are more often phrased separately and realized with a falling contour compared to neg-pos QTs. Since our statistical model includes all factors (polarity, position, verb type) in one model, this reported difference in polarity holds over and above the other factors. What we do not know, however, is how polarity interacts with pragmatic factors. To be able to arrive at an even more complete picture, more prosodically annotated data would be desirable, in particular to be able to also investigate interactions between the factors tested. Furthermore, semantic and pragmatic effects on the prosody of QTs have not yet been systematically studied on the basis of a large data set.

Our data further show an effect of turn-position on the prosodic realization of QTs. At the end of a speaker's turn, QTs are more often realized with a rising intonation, thus inviting a response from the interlocutor; more so than in sentence-final QTs that were not turn-final. These findings shed new light on the role of boundary tones for discourse organization and show that continuation rises (holding the floor) may be limited to declarative sentences. It would be ideal to complement the present corpus data by controlled perception experiments to corroborate our interpretation of these interactional factors. Seen in a larger context, such perception experiments would also allow us to disentangle performance factors from factors that signal functional contrasts.

Authors' address:

Fachbereich Sprachwissenschaft

Universität Konstanz

Fach 186

78457 Konstanz

Germany

nicole.dehe@uni-konstanz.de

bettina.braun@uni-konstanz.de

References

- Baayen, Harald R. 2008. *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Bald, Wolf-Dietrich. 1980. English tag-questions and intonation. In Kuno Schuhmann (ed.), *Anglistentag 1979: Vorträge und Protokolle*, 263–91. Berlin: Technische Universität Berlin.
- Beckman, Mary E. & Gayle Ayers Elam. 1993. Guidelines for ToBI labelling: The Ohio State University Research Foundation. 43 pp.
www.ling.ohio-state.edu/~tobi/ame_tobi/labelling_guide_v3.pdf (17 June 2012)

- Beckman, Mary E. & Janet B. Pierrehumbert. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3, 255–309.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson Education.
- Bing, Janet M. 1985. *Aspects of English prosody*. New York: Garland.
- Boersma, Paul & David Weenink. 2012. Praat: Doing phonetics by computer [Computer program]. Version 5.3.04. www.praat.org.
- Bolinger, Dwight. 1989. *Intonation and its uses: Melody in grammar and discourse*. London: Edward Arnold.
- Cattell, Ray. 1973. Negative transportation and tag questions. *Language* 49(3), 612–39.
- Cooper, William & John Sorensen. 1981. *Fundamental frequency in sentence production*. Heidelberg: Springer.
- Cruttenden, Alan. 1997. *Intonation*, 2nd edn. Cambridge: Cambridge University Press.
- Crystal, David. 1969. *Prosodic systems and intonation in English*. Cambridge: Cambridge University Press.
- Dehé, Nicole. 2009. Clausal parentheticals, intonational phrasing, and prosodic theory. *Journal of Linguistics* 45(3), 569–615.
- Emonds, Joseph. 1976. *A transformational approach to English syntax*. New York: Academic Press.
- Ferreira, Fernanda. 1993. Creation of prosody during sentence production. *Psychological Review* 100, 233–53.
- Gussenhoven, Carlos. 1984. *On the grammar and semantics of sentence accents*. Dordrecht: Foris.
- Gussenhoven, Carlos. 2004. *The phonology of tone and intonation*. Cambridge: Cambridge University Press.
- Gussenhoven, Carlos & Toni Rietveld. 1992. Intonation contours, prosodic structure and preboundary lengthening. *Journal of Phonetics* 20, 283–303.
- Hirschberg, Julia. 2004. Pragmatics and intonation. In Laurence R. Horn & Gregory Ward (eds.), *The handbook of pragmatics*, 515–37. Oxford: Blackwell.
- Huddleston, Rodney. 1970. Two approaches to the analysis of tags. *Journal of Linguistics* 6(2), 215–22.
- Huddleston, Rodney & Geoffrey K. Pullum *et al.* 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Knowles, Gerry. 1991. Prosodic labelling: The problem of tone group boundaries. In Stig Johansson & Anna-Brita Stenström (eds.), *English computer corpora: Selected papers and research guide*, 149–61. Berlin: Mouton de Gruyter.
- Knowles, John. 1980. The tag as a parenthetical. *Studies in Language* 4(3), 379–409.
- Kreiman, J. 1982. Perception of sentence and paragraph boundaries in natural conversation. *Journal of Phonetics* 10, 163–75.
- Krivokapić, Jelena. 2007. Prosodic planning: Effects of phrasal length and complexity on pause duration. *Journal of Phonetics* 35(2), 162–79.
- Ladd, D. Robert. 1981. A first look at the semantics and pragmatics of negative questions and tag questions. *Chicago Linguistic Society (CLS)* 17, 164–71.
- Ladd, D. Robert. 1986. Intonational phrasing: The case for recursive prosodic structure. *Phonology Yearbook* 3, 311–40.
- Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press.
- Ladd, D. Robert. 2008. *Intonational phonology*, 2nd edn. Cambridge: Cambridge University Press.
- Lehiste, Ilse. 1973. Phonetic disambiguation of syntactic ambiguity. *Glossa* 7, 107–22.
- McCawley, James D. 1998. *The syntactic phenomena of English*, 2nd edn. Chicago, IL: University of Chicago Press.

- Nelson, Gerald, Sean Wallis & Bas Aarts. 2002. *Exploring natural language: Working with the British component of the International Corpus of English*. Amsterdam: John Benjamins.
- Nespor, Marina & Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris.
- Potts, Christopher. 2002. The syntax and semantics of *as*-parentheticals. *Natural Language and Linguistic Theory* 20(3), 623–89.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. New York: Longman.
- Rando, Emily. 1980. Intonation in discourse. In Linda R. Waugh & C. H. van Schooneveld (eds.), *The melody of language*, 243–78. Baltimore: University Park Press.
- Reese, Brian & Nicholas Asher. 2008. Prosody and the interpretation of tag questions. In Estela Puig-Waldmüller (ed.), *Proceedings of Sinn und Bedeutung* 11, 448–62. Barcelona: Universitat Pompeu Fabra.
- Sadock, Jerrold. 1974. *Toward a linguistic theory of speech acts*. New York: Academic Press.
- Selkirk, Elisabeth. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology Yearbook* 3, 371–405.
- Selkirk, Elisabeth. 1995. Sentence prosody: Intonation, stress, and phrasing. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 550–69. Oxford: Blackwell.
- Selkirk, Elisabeth. 2009. On clause and intonational phrase in Japanese: The syntactic grounding of prosodic constituent structure. *Gengo Kenkyu* 136, 35–73.
- Selkirk, Elisabeth. 2011. The syntax–phonology interface. In John A. Goldsmith, Jason Riggle & Alan C. L. Yu (eds.), *The handbook of phonological theory*, 2nd edn, 435–84. Oxford: Wiley-Blackwell.
- Silverman, Kim E. A., Mary E. Beckman, John F. Pitrelli, Mari Ostendorf, Colin Wrightman & Patti J. Price. 1992. ToBI: A standard for labeling English prosody. *Proceedings of the 2nd International Conference on Spoken Language Processing (ICSLP)*, Banff, Canada.
- Syrdal, Ann K. & Julia McGory. 2000. Inter-transcriber reliability of ToBI prosodic labeling. *Proceedings of the 6th International Conference on Spoken Language Processing (ICSLP)*, Beijing, China.
- Tottie, Gunnel & Sebastian Hoffmann. 2006. Tag questions in British and American English. *Journal of English Linguistics* 34(4), 283–311.
- Truckenbrodt, Hubert. 2002. Upstep and embedded register levels. *Phonology* 19(1), 77–120.
- Turk, Alice E. & Stefanie Shattuck-Hufnagel. 2007. Multiple targets of phrase-final lengthening in American English words. *Journal of Phonetics* 35(4), 445–72.
- Vaissière, Jacqueline. 1983. Language-independent prosodic features. In Anne Cutler & D. Robert Ladd (eds.), *Prosody: Models and measurements*, 53–66. Berlin: Springer.
- Wichmann, Anne. 2007. Can English tag questions grammaticalise? *Nouveaux cahiers de linguistique française* 28, 349–57.
- Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf & Patti J. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91(3), 1707–17.
- Ziv, Yael. 1985. Parentheticals and functional grammar. In A. Machtelt Bolkestein, Casper Groot & J. Lachlan Mackenzie (eds.), *Syntax and pragmatics in Functional Grammar*, 181–213. Dordrecht: Foris.