

A functional model for the tag question paradigm: The case of invariable tag questions in English and Portuguese



María de los Ángeles Gómez González^{*}, Maria da Purificação Moura Silvano¹

University of Santiago de Compostela, University of Porto, Portugal

Received 5 August 2021; revised 18 January 2022; accepted in revised form 19 January 2022; Available online 28 February 2022

Abstract

While research has mostly focused on the pragmatics of variable tag questions, fewer studies have explored invariable tags, either for their own sake or in contrast with other tag types within and across languages. It will be argued that invariable tag questions are as much part of the tag question system as variable tag questions, and that a unified functional model needs to consider both types in order to compare function-to-form mappings and reveal language variation, as well as the factors motivating their use. This study proposes one such model comprising eight functional types of tag questions, i.e., *informational*, *affective*, *challenging*, *hortatory*, *facilitative*, *focusing*, *phatic* and *regulatory*, in relation to four clusters of grammatical, dialogic, generic and sociolinguistic features. Based on the analysis of 539 invariable tag questions in British English and European Portuguese, results show that the constructions are more frequent and functionally varied in Portuguese (N = 397 vs. 142). In addition, based on statistical analyses, corresponding multi-feature prediction models are obtained for the proposed functional types of invariable tag questions in the two languages under inspection, thereby uncovering novel contributions to the pragmatics of invariable tag questions within the tag question paradigm.

© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Keywords: Tag questions; Invariable question tags; British English; European Portuguese

Abbreviations: TQ, Tag Question; QT, Question Tag; ANC, Anchor; ITQ, Invariable Tag Question; VTQ, Variable Tag Question; BE, British English; EP, European Portuguese; BEITQ, British English Invariable Tag Question; EPITQ, European Portuguese Invariable Tag Question

^{*} Corresponding author at: Department of English and German, University of Santiago de Compostela, Avda. Castelao, s/n. Campus norte, 15782 Santiago de Compostela, Spain.

E-mail address: mdelosangeles.gomez@usc.es (M. de los Ángeles Gómez González).

URL: <http://www.scimitar.es/> (M. de los Ángeles Gómez González).

¹ Address: Faculty of Arts and Humanities of the University of Porto, Centre of Linguistics of the University of Porto, Portugal.

<https://doi.org/10.1016/j.lingua.2022.103255>

0016-7037/© 2022 The Author(s). Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. INTRODUCTION

As the body of literature continues to expand, in English functional research has mostly focused on the pragmatics of variable tag questions (*You've seen Martin's hall haven't you?*) (for overviews see e.g., [Kimps, 2018](#); [Gómez González, 2020](#)). Investigations comparing the use of variable and invariable tag questions (*Well what the hell eh?*) are more scattered ([Holmes, 1983](#); [Biscetti, 2006](#); [Tottie and Hoffmann, 2006](#); [Allerton, 2009](#); [Axelsson, 2011a](#); [Bonsignori, 2013](#); [Kimps et al., 2014b](#); [Gómez González, 2018, 2020](#), among others). Likewise, relatively few and far between are studies focusing on invariable tag questions (e.g., [Cuenca, 1997](#); [Stenström et al., 2002](#); [Biscetti, 2006](#); [Columbus, 2010](#); [Mithun, 2012](#); [Pichler, 2013](#)), attention being generally centered on specific invariable tags such as *eh?*, *ok?* and *no?* ([Bublitz, 1978](#); [Andersen, 1997, 2001](#); [Berland, 1997](#)) or on the increasing invariable use of variable tag questions such as *isn't it?*, *innit?* and *is it?* ([Algeo, 1990](#); [Stenström et al., 2002](#); [Columbus, 2010](#); [Achiri-Taboh, 2015](#); [Criado Peña, 2016](#); among others).

In Portuguese, the literature is sparser as neither variable nor invariable tag questions have been subjected to an in-depth functional description in their own right, but rather they have been studied only in passing alongside other interrogative or polarity-related constructions ([Santos, 2003](#); [Martins, 2006](#); [Barbosa et al., 2020](#), among others). Exceptions include [Cruz-Ferreira's \(1981\)](#) intonational and morpho-syntactic study, as well as some contrastive investigations ([Duarte, 1985](#); [Recksky, 2006](#); [Gómez González, 2014](#); [Carvalho and Kern, 2019](#)). As to the analyses of specific invariable tag questions, they do not abound and are mainly concerned with *não é*, *não* and *pois não* (e.g., [Cruz-Ferreira, 1981](#); [Rodrigues, 1999](#); [Santos, 2003](#); [Martins, 2006, 2013](#)).²

Despite the differences, most prior accounts explain the pragmatics of both variable and invariable tag questions along analogous *interactional* and/or *stance* lines assuming that they function similarly ([Kimps et al., 2014b](#); [Kimps, 2018](#); [Martínez Caro, 2020](#)). Interactionally, by virtue of the polar interrogative patterns, tag questions have been described as *turn-allocation* or *exit* devices with which speakers conduce the structure of conversation towards the response intended and expected from the proposition to which they are appended. As a result, they show a tendency to form adjacency pairs with regard to some information or desired action. From a stance perspective, on the other hand, tag questions have been said to behave as stance modifiers assigning different degrees of commitment on the part of the speaker towards the *epistemic stance* (roughly assessment of the likely truth or accuracy) or the *attitudinal stance* (affective attitude or appreciation) of the proposition they are attached to.

From the above it follows that their dialogic dimension (i.e., whether verbal responses are provided or not) is criterial to determine the functionalities of tag questions. Similarly, their context of use including genre constraints and users' roles has been found to be a motivating factor for the choice of a particular tag type and function over another ([Calnan and Davidson, 1998](#); [Cameron et al., 1988: 86](#); [Kimps et al., 2014a](#); [Kimps, 2018](#); [Cheng and Warren, 2001](#); [Gómez González, 2012, 2014, 2016, 2018](#), among others). In addition, prior socio-linguistic research has found that certain types of tag questions are used more by men (e.g., modal tags) ([Dubois and Crouch, 1975](#); [Lapadat and Seesahai, 1977](#)), whereas others are mostly employed by women (e.g., facilitative tags) ([Lakoff, 1975](#); [Holmes, 1984](#); [Cameron et al., 1988](#); [Eckert and McConnell-Ginet, 2003](#)), which apparently suggests that both males and females use tag questions but in different ways.

Nevertheless, the question remains whether variable and invariable tag questions are always interchangeable. Examples like (1) ([Achiri-Taboh \(2015 : 52\)](#)) and (2) from the languages under inspection show that this is not always the case because, in specific situations, the substitution of a variable tag question by an invariable one (and vice versa) may render ungrammatical or pragmatically different constructions.³

² Additional references on Brazilian Portuguese are [Duarte \(1985\)](#), [Recksky \(2006\)](#), [Carvalho and Kern \(2019\)](#), *inter alios*.

³ Unless otherwise specified, the examples are extracted from ICE-GB and the European Portuguese component of C-ORAL-ROM (see Section 3). Hence, in ICE-GB the citation <S1b-062 #137: B> corresponds to mode (S = spoken), section (1), subsection (b), text number (062), turn (#137), subtext (1) and speaker (B). Turns are given without punctuation and the only prosodic detail provided in this corpus is the marking of pauses <,> and longer pauses <,>. Likewise, in the Portuguese examples the annotation <CORALRpnatbu01> stands for language ('p'-Portuguese), the text field ('nat'-natural context) and subfield ('bu'-business), as well as its code number (01). Additional conventions for prosodic tagging include: // (conclusive prosodic break), / (non-conclusive break) and ? (optional interrogative mark).

-
- | | | |
|----|----|---|
| 1. | a. | It's getting warmer, isn't it/*okay? |
| | b. | Leave the cat alone, will you/*right? |
| | c. | How quiet it was in there, wasn't it/*okay? |
| | d. | Let's go home now, shall we/*right? |
| 2. | a. | <Foste tu > que me mandaste / < não foste? >/ # okay ? <CORALRpfamnn13>
'It was you who sent me, weren't you/*okay?' |
| | b. | Porque finalmente era / muito aberto / < não era >/ # está bem? <CORALRpfamnn23>
'Because it was finally open, wasn't it/*is it ok?' |
| | c. | / eu vou passar a sua mãe / que quer falar consigo / < está bem >/* não vou? <CORALR ptelpv08>
'I'll pass to your mother, who wants to talk to you, ok?' |
-

It is here argued that variable and invariable tag question alternations are not felicitous, or at least are not pragmatically equivalent, if they trigger a change in the interpretation of the construction. Accordingly, the grammaticality or acceptability of the tag question alternations in (1) and (2) is attributed to the change in the functional type of the construction triggered by replacing the question tag, from (*confirmatory*) epistemic to *challenging* or (*phatic*) *textual*. This is at odds either with the type of stance encoded in the constituent which the tag is appended to – in the asterisked examples in (1) and (2) –, or with the context of situation when the alternations require different settings to be adequately used. The latter is the case of the asterisked examples in (1a) to (1c), which unlike their counterparts are felicitous only if what is said is interpreted as an annoyed comment on something that has been resisted by the addressee. This evidence underpins the necessity for a fine-grained functional analysis of the tag question paradigm within and across languages to reveal language variation, as well as their conditions of use.

This paper homes in on this gap by applying a hybrid theoretical and empirical procedure. It consists in presenting a functional framework for the analysis of the tag question paradigm to be empirically validated on the basis of invariable tag questions across two languages. An in-depth characterization of variable and invariable tag questions in English and Portuguese is accordingly first offered (Section 2.1). The comparison is based on these two varieties not only because invariable tag questions have been neglected in the literature in both languages, but also because they represent different families and so linguistic variation is likely to emerge from their comparison. Secondly, a unified functional model for the tag question paradigm is proposed distinguishing eight functions, i.e., *informational*, *affective*, *challenging*, *hortatory*, *facilitative*, *focusing*, *phatic* and *regulatory*, focusing on invariable tag questions (Section 2.2). After the pertinent methodological considerations (Section 3), the descriptive and crosslinguistic adequacy of this model is attested on the basis of 539 invariable tag questions taken from spoken data in British English and European Portuguese concerning fifteen variables of analysis (Appendix A) (Section 4). The overall purpose of the study is to unveil the “bundles” of features that characterize each and every proposed functional type of invariable tag questions. Specifically, relying on our own and prior work on tag questions, five main research questions are addressed, after which the study concludes with a summary of the main findings:

RQ # 1. What are the frequencies of invariable tag questions in British English and European Portuguese, and what are the pragmatic motivations underpinning their use? (Section 4.1)

RQ # 2. What are the grammatical characteristics of the proposed functional types of invariable tag questions? Do they show any similarities across the two languages despite belonging to different language families? (Section 4.2)

RQ # 3. What are their dialogic and generic features? (Section 4.3)

RQ # 4. Can any aspect of sociolinguistic variation be determined regarding their use? (Section 4.4)

RQ # 5. Can a model of predicted associations be proposed based on the statistically significant results obtained from RQ # 1 to RQ # 4? (Section 4.5)

2. THE TAG QUESTION PARADIGM WITHIN A FUNCTIONAL FRAMEWORK

2.1. Variable and invariable tag questions in English and Portuguese

English *tag questions* (Jespersen, 1924: 323) have been characterized as utterances or *anchors* (ANCs) with interrogative tags or *question tags* (QTs) (in bold type) that are consecutively produced by the same speaker (Aijmer, 1979; Huddleston and Pullum, 2002: 891; Axelsson, 2011a, b; Kimps, 2018; Gómez González, 2012, 2014, 2016, 2018, 2020,

inter alios). As will be shown, the construction results from a modification of the ANC by the QT (Kimps, 2007: 270; Gómez González, 2018: 113) or from a combination of the two and their corresponding characteristics (Hudson, 1975; Nässlin, 1984: 26). ANCs are normally realized by declarative matrix clauses – (3a), (3b), (4a), (4b) –, but they may also be interrogatives – (3c), (4c) –, imperatives – (3d), (4d) – or exclamatives – (3e), (4e). QTs, in turn, are generally appended to and separated off from the ANCs on which they are grounded. They may be variable / variant or invariable / invariant, turning the construction into *variable / variant* (VTQs) – (3) – or *invariable / invariant* tag questions (ITQs) – (4) –, respectively (for overviews see e.g., Allerton, 2009; Kimps, 2018; Gómez González, 2018).

-
- (3) a. You've seen Martin's hall **haven't you?** < ICE-GB:S1A-073 #4:1:B>
 b. He doesn't seem to care though **does he?** which isn't such a bad thing < ICE-GB:S1A-041 #284:1:B>
 c. Isn't it true to say that if you get the get the design of the fermentor wrong <, > the costing wrong then the rest of the process is a little bit academic **isn't it?** <, > <ICE-GB:S1B-020 #48:1:A>
 d. Let's stop for the moment **shall we?** < ICE-GB:S1A-001 #051>
 e. Oh what a mess everywhere **isn't it?** (Roesle, 2001: 52)
-

- (4) a. You know tight skin jeans that went choo choo **don't you think?** <, > <ICE-GB:S1A-085 #011:1:B>
 b. I'll give you a ring next time I'm down and we'll try and meet **OK?** <ICE-GB:S1A-098 #131:1:A>
 c. Is Bim at the Slade now **or not?** < ICE-GB:S1A-015 #072:1:B>
 d. Dani come on you're not doing anything start talking **yeah?** <, > <ICE-GB:S1A-038 #011:1:A>
 e. Well what the hell **eh?** <, > <ICE-GB:S1A-039 #211:1:A>
-

Formally, VTQs – (3) – contain a grammatically dependent QT (e.g. *haven't you?*, *does he?*, *isn't it?*, *shall we?*) that is structurally and lexically conditioned by their ANCs: (i) reproducing, if explicit, the same operator or auxiliary with identical tense, number, and person; (ii) repeating, if present, the Subject (S) of the ANC, usually a personal pronoun agreeing in number, person and gender; and (iii) if negative, with the operator and enclitic negator (*n't*) before the S (3c) or, less commonly, including *not* after S. VQTs usually reverse the polarity across ANC and QT, either (+ / –) – (3a), (3e) – or (– / +) – (3b). Less frequently, polarity may remain constant, either affirmative (+ / +) – (3d) – or negative (– / –) – (3c). Prosodically, VTQs (especially the reversed type) tend to have a nuclear accent, be uttered with a falling tone, and be preceded and followed by phrasal prosodic boundaries in turn-final, sentence-final or phrase-final position, typically obtaining a confirmatory reply (for overviews see e.g., Tottie and Hoffmann, 2006; Kim and Ahn, 2008; Axelsson, 2011a, 2011b; Kimps, 2018; Gómez González, 2018, 2020).

In contrast, ITQs – (4) – may be triggered by invariable / invariant lexical QTs (e.g. *yes?* / *yeah?*, *ok* / *okay?* – (4b), (4d) – (Quirk et al., 1985: 814, 1115), intonational QTs (*eh?*, *hunh?*) – (4e) – (Bublitz, 1979; Holmes, 1983), reduced *non-concordant yes/no* questions QTs (e.g. (*is that*) *right* (and variations), *isn't that so*, *don't you think*, *wouldn't you say*) – (4a) – (Quirk et al., 1985: 814, 1115), or *disjunctive* QTs like *or not* and *or what* – (4c) – (Govindan and Pillai, 2009). IQTs stay the same regardless of the structure of the ANC, which are generally declarative clauses as in VQT. IQTs may also keep polarity constant – (4d) – or they may reverse it – (4a), (4c). But, unlike their variable counterparts, additional patterns emerge as some IQTs (e.g., *eh?*) – (4e) – have no marks of polarity and therefore are *polarity-neutral*. Furthermore, ITQs generally show a higher tendency to display constant positive polarity patterns and to obtain no response from addressees, among other differences (see e.g., Quirk et al., 1985: 814; Norrick, 1995; Biber et al., 1999: 210 ; Andersen, 2001; Stenström et al., 2002; Columbus, 2010; Axelsson, 2011a; Torgersen et al., 2011; Pichler, 2013; Gómez González, 2018, 2020; *inter alios*).

Turning to Portuguese, since it is a null-Subject language in which pronominal Subjects are not generally lexically realized, IQTs show more variation lexically and phonetically as different forms are used depending on whether they are accented or deaccented (Cruz-Ferreira, 1981; Martins, 2006; Moniz et al., 2011). In addition, we distinguish four types of TQs depending on the degree of dependency of QTs upon their ANCs: *variable*, *partially variable*, *partially invariable* and *invariable*, as illustrated in Fig. 1 (Martins, 2006; Gómez González, 2014).

Constructions like (5) and (6) below occupy opposite extremes as ITQs and VTQs, respectively. In (5), *não é verdade?*, *certo?* and polarity-neutral *eh?* are regarded to trigger ITQs because they stay the same regardless of the morpho-syntactic features of the ANC (tense, verbal agreement and polarity); whereas *não sabe?* in (6) is classified as variable because it changes in accordance with such features. In contrast, examples like (7) to (9) show in-between degrees of dependency or variability. *Não foi?* – (7) – is a partially variable TQ because it reproduces the tense

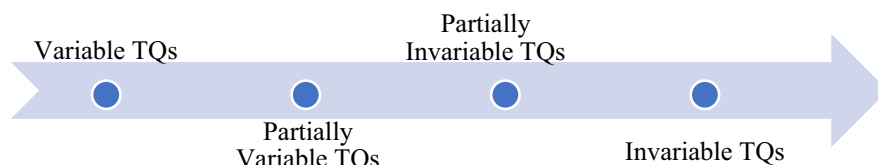


Fig. 1. TQ variability in Portuguese.

and verb agreement of the ANC, but it uses the verb *ser* 'be' instead of repeating that of the ANC and is restricted to reversed polarity patterns. Examples like *não é?* (8) and *pois não?* (9), in turn, are illustrations of the partially invariable type because they stay the same regardless of the tense and verbal agreement features of their ANCs, although they have a more restricted distribution than IQTs. In their invariant use, *não é?* is biased towards reversed polarity patterns (+ / -) while *pois não?* occurs only in constant negative (- / -) patterns.⁴

-
- (5) a. *é o big brother que está lá em cima / não é verdade?* <CORALRpmedin03>
 'it's the big brother who is up there, isn't that right?'
 b. *eu começo aqui a fazer uma ordenação // <|certo|> ? e depois esta começa logo aqui a seguir // e / &eh* <CORALRpntate02>
 'Here I start to put things in order right? and then it starts right after eh'
 c. *não era político // <hã ? nem era militar // <CORALRpmedts02>*
 'he wasn't a politician eh? not a soldier either'
-

-
- (6) *sabe onde é o liceu Filipa de Vilhena/ não sabe?* <CORALRppubdl03>
 'you know where Filipa de Vilhena high school is, don't you?'
 (7) *E ele estudou / ele estudou/ fora / não foi?* <CORALRpfammn15>
 'And he studied, he studied abroad, didn't he?'
 (8) *Em princípio não improvisarão / não é ?* <CORALRpfamcv10>
 'In principle they will not improvise, right?'
 (9) *E / não há dúvidas / pois não?* <CORALRpfammn01>
 'And there are no doubts, are there?'
-

Both VTQs and ITQs may deviate from these general tendencies. Focusing on the latter, one such case is *isn't it?* – (10) –, the most common variable tag question in Standard British English and American English (Tottie and Hoffmann, 2006; Gómez González, 2014, 2018), which is pervasively used as an ITQ in several (colloquial) varieties of English (i.e., Welsh, Singapore, Malaysian, Indian, Hong Kong or Cameroonian English) (Achiri-Taboh, 2015; Takahashi, 2014; Criado Peña, 2016). A similar phenomenon occurs in Portuguese with the variable QT *não é?*. It may be used as a partially invariable QT – (8) – and as an IQT that “is able to replace all other tags except the disjunctive ones” and “carries no meaning of tense” even in the standard varieties, as explained by Cruz-Ferreira (1981: 346). In addition, tag questions may be added to phrases or to incomplete or *truncated* clausal ANCs – (11) –, or they may also occur medially – (12) – or be stacked in the QT slot, in which case they either stick to the one type – (13) –, or combine both – (12a) (Gómez González, 2014, 2018, 2020). The sequencing of QTs offers a research window on the discourse-functional structure of the right clause periphery within and across languages, mirroring the increasing number of investigations on discourse marker organization in the left clause periphery (Fraser, 2013; Tagliamonte, 2016). Such an analysis, however, lies beyond the scope of this study.

⁴ For further discussion of the issue of tag question variability from a cross-linguistic perspective, see Tottie and Hoffmann (2006), Allerton (2009), Axelsson (2011b), Mithun (2012), and Hoffmann et al. (2014), among others.

-
- (10) John doesn't eat fish, **isn't it?** (Achiri-Taboh, 2015: 51)
- (11) a. Shelf on the top front room **OK?** < ICE-GB:S1B-073 #140:1:C >
 b. *JOS: [<] < o factor > cultural / <não acha> ? CORALRppubdl10>
 'JOS: the cultural factor, don't you think?'
- (12) a. But everybody talks about them **you see? don't they?** as being so marvellous <, > < ICE-GB:S1A-016 #203:1:D >
 b. Se não tivesse havido / o dinheiro / inicial / posto / pelo programa ciência / de / de / da / da &j / da JNICT da altura / <não é? / da junta> / de investigação científica // <CORALR- pfammn17>
 'If it hadn't been for the initial money put in by the science program of the JNICT at the time, hadn't it? from the council of scientific research'
- (13) B. Let's look at the women. So we can discuss it yeah **OK yeah right** < ICE- GB:S1A-090 #131:1:B>
 C. Ok brilliant
-

2.2. A functional model for the tag question paradigm

This study proposes a unified functional model – Table 1 – for the description of the tag question paradigm within and across languages. Extending prior work that used this framework to explore mostly VTQs across languages (Gómez González, 2014, 2018, 2020), the model is now applied to compare and contrast ITQs in British English (BEITQs) and European Portuguese (EPITQs) according to four kinds of stance (*epistemic*, *deontic*, *attitudinal*, and *textual*) and eight functionalities (*informational*, *affective*, *challenging*, *hortatory*, *facilitative*, *focusing*, *phatic* and *regulatory*), with corresponding descriptions at the locutionary, illocutionary and perlocutionary levels.⁵

The first stance type corresponds to statement-question ITQs used to negotiate shared knowledge, comprising such examples as (4a), (4c), (5a) or (5c) above categorized as epistemic and information/verification-seeking that are in many cases overlapping. The category includes ITQs that expect the supply of additional information or verification, as well as those that trigger a confirmatory response regarding information known to both speaker and addressee, assuming that these uses often coalesce in natural discourse.⁶ In this case speakers position themselves along an epistemic or knowledge scale including different degrees of certainty, likelihood and evidentiality (grounds for speakers' judgments) (Biber et al., 1999, 2018; Boye, 2012).

Deontic or action-seeking ITQs are used as softening strategies to get the addressee to do something. They express deontic stance or the speakers' positions towards three main parameters: necessity / obligation, permission / possibility / ability and causation / effort (Lyons, 1977), as represented in the above examples concerning a desired action. Two subtypes of deontic ITQs are recognized: *hortatory* and *facilitative*, which generally are indexed by expressions of willingness or commitment (*ok?*, *eh?*), as well as affirmative polarity markers (*yeah?*, *right?*, *está bem?*), and mostly have the perlocutionary effect of triggering the desired actions. As hortatory devices, TQs range from *demanding* as in commands and requests – (4b), (13) – to *offering* and *invitations* – (14). A variant of the latter, facilitative are predominantly used as positive politeness strategies to invite the interlocutor(s) to take the discourse floor and continue talking, as shown in (4d) and (15).

-
- (14) a. B: See how the gig goes **eh?** < ICE-GB:S1A-099 #348:2:B>
 b. /depois a gente vê isso> / <|está bem|> ? <CORALRptelpv11>
 'then, we will see that, ok?'
-

⁵ Although ITQs are potentially multifunctional and consequently may target the joint expression of more than one meaning at a time (Fischer, 2006), we assume they can be ascribed a predominant or primary function in a specific discourse context. This decision, however, ultimately, depends on the speaker's or analyst's interpretation.

⁶ Other investigations (Brazil, 2008; Axelsson, 2011a, 2014; Kimps, 2016) distinguish three different subtypes of information-seeking TQs depending on the event-type encoded: *A-event* (known to A, but not to B), *B-event* (known to B, but not to A), and *AB-event* (known to both A and B) (Labov, 1972; Labov and Fanshel, 1977).

- (15) *JOS: [<] < o factor > cultural / < **não acha** > ? <CORALRppubdl10>
 'JOS: the cultural factor, don't you think?'
 *NAZ: [<] < cultural > também // sim // também o factor cultural / acho que sim // também // isso
 'Naz: cultural, also, yes, also the cultural factor, I think that yes, also that.'

Within the attitudinal category, *affective* ITQs show speakers' attitude to propositions, conveying subjective evaluation / appreciation, frustration, and so on (Biber et al., 1999, 2018; Hyland, 2005a). They are often indexed by means of an adjective or some sort of (usually intensified) adjectival phrase (*tão lindo* 'so nice') or evaluative expression (*problem*), as in (16) below. *Challenging* ITQs, on the other hand, are generally used as confrontational strategies to boost the force of a negative, threatening, or aggressive speech act. Within this group are also included ITQs expressing *mirativity* or surprise, that is, reactions of the "unprepared mind" (Andersen, 1998: 308–311) towards an event "somehow not consonant with the current state of mind of the speaker" (Aksu-Koç and Slobin, 1986: 160), mostly marking a disalignment with the co-participant or a perceived incongruence with the situation (Simon-Vandenberg and Aijmer, 2007; Kimps, 2007, 2018). This occurs in (4e) above, where the "unprepared mind" expresses an exclamative ANC with an evaluation noun (*hell*), as well as in (17) below, where a scornful criticism is made about the seaworthiness of an old vessel (*vossa caravela*): the counter-expectation (that it can still sail despite its many years) triggers a humorous effect (see also e.g., Algeo, 1990, 2006; Holmes, 1995; Biscetti, 2006; Tottie and Hoffmann, 2006).

- (16) a. Mummy's always had this Vienna problem **OK?** <,> ICE-GB:S1A-032 #264:2:A>
 b. Era um tempo tão lindo // **não é /?** <CORALRpfamd103>
 'It was such a nice time, wasn't it?'

Table 1
 A functional framework for BEITQs and EPITQs.

Focus on	Stance Types	Functional Types	Illocutionary level	Perlocutionary level	
Interactional	Addressee	Epistemic	Informational	Statement- Question Blends	Answer Verification Confirmation Acknowledgement vs. Disclaimer Refutation Contradiction
		Deontic	Hortatory Facilitative	Order Command Suggestion Request Invitation	Undertaking vs. Denial
	Speaker	Attitudinal	Affective Challenging	Evaluation Threat Joke	Agreement Response Appreciation vs. Disagreement Rejection
					No verbal response: Backchannel feedback (not evident) No response or paralinguistic features
Interactive	Interaction	Textual	Focusing Regulatory Phatic	Follow-ups Repairs	Confirmation Verification Backchannels vs. Refutation

-
- (17) Desde que a vossa caravela da / hhh / dos quinhentos anos não navegou / **não é** / ? <CORALRpfdm104>
'Since your five hundred caravel didn't sail, right?'
-

Finally, *textual* ITQs are used as textual stance markers to organize the conversation in reflection of the speakers' line of reasoning or topic management strategies (Tomaselli and Gatt, 2015). This functional type is displayed in (12) above illustrating the *focusing* use (Carter and McCarthy, 2006; Gómez González, 2012, 2014, 2018, 2020). The insertion of the IQTs within their respective ANCs focuses attention on both the pre-tag and post-tag slots that are realized in separate intonation units keeping an informational balance between them as both receive focal prominence. In all these cases, a shared thematic perspective is intended to be acknowledged by the addressee in the pre-tag slot, namely the fact everybody talks in (12a), and the science program of the JNICT in (12b). At the same time, the post-tag constituents are placed in rhematic position within the ITQ and therefore receive end focus prominence, expressing what is said about them (*being so marvellous*) or specifying the funding institution's name (*da junta de investigação científica*).

Also included in the textual category are ITQs fulfilling a *phatic* function for controlling contact – (18) – in the sense that they are used to maintain, verify or follow-up the communication channel and understanding between the participants (Cruz Ferreira, 1981; Holmes, 2001; Gómez González, 2012, 2014, 2018, 2020). The examples in (19), in turn, are assigned to the *regulatory* or "narrative" subtype (Columbus, 2010; Gómez González, 2012, 2014, 2018, 2020). In this type, ITQs help organize the processing of information including topic organization (closing, emphasizing), thematic corrections, false starts, "performance errors" or such idiosyncratic uses as repeating the same TQ within a turn and throughout a conversation.

-
- (18) a. So this F X tends to infinity **right?** <,>
b. ORQ: fomos a um restaurante brasileiro // <<|está bem|>> ? <CORALRpptelpv08>
'we went to a Brazilian restaurant, right?'
-

-
- (19) a. I mean just to give you a sort of swift example supposing uhm you 've got a chain of gas stations a chain of petrol stations uhm and they have one independent competitor who sets up in business selling low petrol **OK?**
b. GRA: Dava então muitas aulas / só / a pessoas / já formadas // normalmente médicos / e engenheiros // <não é> ? <CORALRpfdm03>
'GRA: At the time you taught many classes only to graduates normally doctors and engineers, right?'
-

The four kinds of stance and eight functional types of ITQs described so far are grouped into two more general categories: *interactive* or *interactional* (Thompson, 2001; Hyland, 2005a, b; among others). In their interactive use, ITQs are resorted to as (focusing, phatic or regulatory) textual cues that guide addressees through the interactions. In their interactional capacity, ITQs are (epistemic, deontic or attitudinal) discourse strategies used to involve the co-participants collaboratively in the development of the discourse, as will be further detailed in Section 4. But let us first clarify some methodological issues.

3. DATA AND METHODOLOGY

Linking up with prior research (Gómez González, 2014, 2018), the data were extracted from two available corpora: the spoken component of the *International Corpus of English-Great Britain – Release 2* (ICE-GB) (Nelson et al., 2002), composed of 637,562 words and 300 texts, and the European Portuguese subcorpus of the *Integrated Reference Corpora for Spoken Romance Languages* (C-ORAL-ROM) (Cresti and Moneglia, 2005), consisting of 384,123 words and 152 texts. For the retrieval, the search utility programs ICE-CUP III (ICE-GB) and Contextes (C-ORAL-ROM) were used. We ran lexical searches for the IQTs listed and aligned as translation equivalents in Table 2, based on prior literature on the constructions in the two languages (BEIQTs and EPIQTs) (Bublitz, 1979; Holmes, 1983; Quirk

Table 2
 IQTs under inspection in this study.

BEIQTs	EPIQTs
1. all right?	1. está bem?
2. am I right?	2. está certo?
3. right?	3. estou certo?
	4. (estás de) acordo
	5. (está) correto?
	6. certo?
	7. verdade?
4. is that right?	8. está certo?
5. isn't that right?	9. não está certo?
6. isn't that so?	10. não é verdade?
	11. não é assim?
7. don't you think?	12. não acha(s)?
	13. não te parece?
8. eh?	14. eh?
	15. hã?
	16. hhh?
9. okay?	17. okay? / ok?
10. is it ok?	18. está okay?
11. or not?	19. ou não?
12. or what?	20. ou quê?
13. see?	21. vê?
14. you see?	22. está(s) a ver?
15. yeah?	23. sim?
16. innit? / VTQ	24. não é? / né? (contracted form)
17. no? / VTQ	25. não?
VTQ	26. não já?
	27. pois não?

et al., 1985; Andersen, 2001; Columbus, 2010; Gómez González, 2014, 2018; Cruz-Ferreira, 1981; Rodrigues, 1999; Santos, 2003; Martins, 2009, 2013).⁷

We had to prune and code the data manually as neither of the two corpora is tagged for ITQs. So, the next step involved determining if the tokens retrieved (cf. Table 2) followed the criteria to be considered BEIQTs and EPIQTs, both structurally (QT invariance irrespectively of ANC features) and prosodically (phrasing and intonation contour).⁸ This task was especially daunting in Portuguese because, among other reasons, variable and partially variable – (7) – QTs with *ser?* had to be discarded, whereas instances of invariable and partially invariable QTs with *não é?* – (8) – had to be included in the sample. We tagged and annotated the selected BEITQs and EPITQs on the basis of the sixteen variables specified in Appendix A, as illustrated in Table 3 comparing the analyses of the ITQs in (20).

- (20) a. B. I can put there my mother's **right?** <ICE-GB:S1B-080 #238:1:B>
 A. Yeah, OK, and initials
 b. TER: ou / será que / &ah / quando nós estamos a crescer/ as nossas células são substituídas muito mais rapidamente / <|certo|>? <CORALRpmedsc01>
 'or it is that when we are growing up, our cells are replaced much more quickly, right?'
 CON: não // numa fase de crescimento / o ritmo de divisão das células / é muito menor
 'no// in a growth phase / cells' division rate is much slower'

⁷ It will be observed that in Table 2 BEIQTs have several translation equivalent EPIQTs, while established EPIQTs like *não?*, *não é?*, *não já?* and *pois não?* may be translated by equivalent BEIQTs (*innit?*, *no?*) or by variable questions tags (VQTs) which would change in accordance with their ANCs. However, in the English translations of the EPIQTs offered in this paper, whenever possible, TQs will be kept invariable.

⁸ Phrasing (whether integrated within a syntactic unit or turn, or separated from it) and intonation contour (falling or rising) helped us determine the structural configuration of invariable question tags. However, a systematic analysis of the prosody of invariable tag questions lied beyond the scope of this work. Details about the prosody of question tags and tag questions can be found in Cruz-Ferreira (1981), Moniz et al. (2016), Dehé and Braun (2013), Kimps (2018) or Gómez González (2020), among others.

Table 3
Exemplification of empirical analysis.

Variables and values	(20a)	(20b)
A. Dependent variables		
1. Function	information	phatic*
B. Independent variables		
Bi. Grammatical		
2. IQT type	<i>right?</i>	<i>certo?</i>
3. IQT position		final
4. Meta-polarity		constant
5. Polarity		positive-positive
6. Mood (ANC)		declarative
7. Truncated ANC		no
8. Clause type (ANC)	simple	
9. Type of Subject (ANC)	pronoun	complex-sub
10. Type of subject (IQT)		noun
11. IQT tense		not applicable
12. ANC tense		not applicable
		present
Bii. Dialogic		
13. Response		other reply
Biii. Generic		
14. Genre		public dialogic discourse
Biv. Sociolinguistic		
15. Sex		female
16. Age	18–25	unknown

*This example has been analyzed as phatic because it is used to verify the correct understanding. Alternatively, it could also be classified as informational if the emphasis is laid on the provision of information about cell reproduction.

To avoid data dispersion in the genre variable, text types were organized according to three parameters that are used in both corpora to classify texts, i.e., monologic discourse (e.g., political speeches, legal presentations, conferences, or weather forecasts), private/informal dialogic discourse (e.g., face-to-face conversations or telephone conversations) and public/formal dialogic discourse (e.g., interviews, political debates or broadcast discussions), as displayed in Table 4. This allowed us to compare ITQ distribution across the private/informal-public/formal and dialogic-monologic divides in the two languages.

Each author analyzed the language sample she felt most confident with, but the most challenging cases were discussed and agreed upon jointly. To ensure that our annotations were consistent and reliable, we fully annotated 30 instances of ITQs in parallel. As a measure of reliability in the annotation, we used Cohen's Kappa (Siegel and Castellan, 1988) for estimating the agreement value. We calculated unweighted kappa values, measuring categorical agreement between the two judges for each of the sixteen labelling variables, which involved two to eleven options (e.g., two for IQT position vs. eleven for ANC tense). In all but one cases, Kappa values were above 0.90, which is considered to be almost perfect agreement (Landis and Koch, 1977). The only value below 0.90 was 0.80 for Function, which is probably the most subjective of the categories due to the multifunctionality of ITQs, as already noted (cf. fn. 4). The agreement is nonetheless also substantial for this category. We thus feel that these are very high levels of agreement, rendering the annotations valid and reproducible.

Finally, to place our results on a firmer footing, statistical analyses were performed in three stages with *R Statistical Software* (R Core Team, 2020) using the packages *vcd* (Meyer et al., 2020) and *Goodman and Kruskal's tau measure* (Agresti, 2012; Pearson, 2020) for estimating association statistics, as well as *Boruta* package for feature selection (Kursa and Rudnicki, 2010a, 2010b). First, in order to answer RQ # 1 to RQ # 4 presented in the introduction, descriptive statistics with percentages and raw frequencies for BEIQTs and EPIQTs were obtained, as well as whenever necessary and possible, normalized frequencies per ten thousand words (nfpttw). In addition, to address RQ # 5 and also to avoid problems derived from cells with low expected counts, χ^2 tests were first used with simulated *p*-values (of expected frequencies) in contingency tables cells in order to evaluate the association between categorical variables. The selected *alpha* (α) was set to 0.05 so that equal or lower *p*-values were considered as significant results. To control the false discovery rate (expected proportion of false discoveries among rejected null hypothesis) *p*-values were then adjusted

Table 4
Classification of tex types in ICE-GB vs C-Oral-Rom-Portuguese.

Genre	ICE-GB text types - codes	C-ORAL-ROM text types - codes
private/informal dialogic discourse	face-to-face conversations - S1A-001 to 090 telephone conversations - S1A-091 to 100	private/family conversations – pfamcv01 to 12 private/family/ dialogues – pfamd101 to 31 private telephone conversations – ptelef01 to 17
public/formal dialogic discourse	classroom lessons - S1B-001 to 020 broadcast discussions - S1B-021 to 040 broadcast interviews - S1B-041 to 050 parliamentary debates - S1B-051 to 060 legal cross-examinations - S1B-061 to 070 business transactions - S1B-071 to 080	public/formal dialogues – pubdl01 to 11 interviews – medin01 to 05 talk shows – medts01 to 03 political debates – pnatpd01 to 05 professional explanation – pnatpe01 business conversation – pnatbu01 to 03 law conversation - pnatla01
monologic discourse	spontaneous commentaries – S2A-001 to 020 unscripted speeches - S2A-021 to 050 demonstrations - S2A-051 to 060 legal presentations - S2A-061 to 070 broadcast news - S2B-001 to 020 broadcast talks - S2B-021 to 040 non-broadcast speeches - S2B-041 to 050	family/private monologues – fmmn1 to 24 public/formal monologues – pubmn01 to 07 meteo/weather forecast – medmt01 to 10 news – mednw01 reportages – medrp01 to 02 scientific reports – medsc01 to 02 sport – medsp01 to 03 political speech – pnatps01 to 03 preaching - pnatpr01 to 02 conferences - pnatco01 to 03 professional explanation – pnatpe02 law monologue – pnatla02 teaching - pnatte01 to 03

using [Benjamini and Hochberg \(1995\)](#) method. The null hypothesis of the test was independence (no association) between the two categorical variables. In a second stage, *contingency coefficient* (CC) and Goodman and Kruskal's tau measures were used to evaluate the association strength between variables: CC (estimated from chi-square and sample size) helped us determine whether two variables were independent or dependent of each other (the closer to one, the greater the association between variables). Goodman and Kruskal's tau measure, on the other hand, allowed us to measure asymmetric associations assuming that the association between variables x and y does not have to be the same as that existing between y and x . The closer the value is to one, the greater the association is between variables. Two types of measures were obtained: *tauxy* $\tau(x, y)$ or *forward associations*, and *tauxy* $\tau(y, x)$ or *backward associations*, where x was our response variable (Function) and y represented each of the independent variables for which association with x was estimated (IQT type, IQT position, etc.). We were interested in finding out whether it was Function that determined the features of ITQs across the independent variables, if forward associations were obtained $\tau(x, y)$; or whether it was the values of the independent variables that predicted the function of ITQs, in which case backward associations would emerge $\tau(y, x)$. Lastly, the Boruta algorithm was employed to identify the most relevant features of the proposed functional types in order to obtain a model of their predicted associations for BEITQs and EPITQs.

4. QUANTITATIVE ANALYSIS: DISCUSSION OF FINDINGS

The descriptive and crosslinguistic adequacy of the unified functional model for the tag question paradigm presented in [Section 2.2](#) is attested here on the basis of 539 invariable tag questions retrieved from the spoken component of ICE-GB and the European Portuguese subcorpus of C-ORAL-ROM concerning sixteen variables of analysis (Appendix A). The overall purpose of the study is to unveil the “bundles” of features that characterize each and every functional type of invariable tag questions posited in this study. Specifically, what follows reveals the findings of the five research questions presented in the Introduction taking *Function* as the observed variable.

4.1. RQ # 1: Frequencies of BEITQs and EPITQs and their functional types

[Table 5](#) shows that in our sample EPITQs are not only five times as frequent as BEITQs ($N = 397$ (10.3 nfptw) vs. $N = 142$ (2.2 nfptt)), but also more functionally varied attesting the eight functionalities considered in the framework ([Table 1](#)).

Table 5

Functional types across BEITQs and EPITQs.

Rank	BEITQs N = 142	EPITQs N = 397
1	informational 75 (52.8%)	regulatory 162 (40.8%)
2	attitudinal 45 (31.7%)	informational 91 (22.9%)
3	focusing 12 (8.5%)	attitudinal 67 (16.9%)
4	hortatory 5 (3.5%)	phatic 51 (12.8%)
5	phatic 4 (2.8%)	focusing 11 (2.8%)
6	regulatory 1 (0.7%)	hortatory 11 (2.8%)
7	challenging (–)	facilitative 3 (0.8%)
8	facilitative (–)	challenging 1 (0.3%)

BEITQs register only six with no tokens of the challenging and facilitative uses. Furthermore, the two languages report different functional rankings. The numbers in Table 5 reveal that EPITQs tend to be used as textual devices (56.4%) (especially regulatory, phatic or focusing) or, to a lesser degree, as statement-question blends to share information (22.9%) or as (dis)alignment strategies to negotiate attitudinal stance (16.9%). In contrast, BEITQs are mostly used as informational (52.8%) or attitudinal strategies (31.7%), the textual (mostly focusing) usage representing only 12%. Deontic ITQs (hortatory and facilitative) record the lowest frequency across the two languages (lower than 4%).

These results complement those reported in previous investigations (e.g., Gómez González, 2014, 2018; McGregor, 1995, 1997; Axelsson, 2011b; Kimps, 2018). Interestingly enough, the same functional scales were obtained for the use of VTQs in the two language varieties under inspection (Gómez González, 2014: 21). It can thus be concluded that in (European) Portuguese, both ITQs and VTQs are predominantly used as interactive or textual strategies (Gómez González, 2014), whereas in (British) English both types of TQs are mostly used in their interactional capacities (epistemic and attitudinal) (McGregor, 1995, 1997; Axelsson, 2011b; Gómez González, 2014; Kimps, 2018), even if BEITQs exhibit a higher incidence of textual usages than the variable type (Gómez González, 2018). This finding, however, not only seems to question previous claims that only a minority of TQs are really information-seeking (Kimps, 2018). This usage ranks first and second in the BEITQs and EPITQs in our sample, respectively, as well as in their variable counterparts according to previous studies (Gómez González, 2014). This result also underscores the importance of the textual capacity of TQs, whether variable or invariable. This aspect has been underexplored in the English literature, but it has received ample attention in studies examining TQs in Romance languages (Cortés and Camacho, 2005; Pons Bordería, 1998; Portolés, 1998; Rodríguez Muñoz, 2009).

In addition, a possible explanation for the scarcity of deontic (hortatory and facilitative) IQTs in the two languages is that speakers generally avoid the face threatening force of expressions with obligation meaning (Biber et al., 1999: 489), and, if they do use them, they tend to resort to other grammatical devices (modal auxiliaries) as deontic markers. Likewise, the informality of the conversations could explain why deontic IQTs may prove unnecessary or inappropriate.

Turning to the association between function and IQT type, Table 6 reveals that, out of the seventeen types of BEITQs, only thirteen were recorded in the sample as there was no record of *isn't that so?*, *is it ok?*, *innit?* and *no?* This result confirms the findings of prior investigations attributing the IQT use of *innit?* and *no?* to other varieties of English (Algeo, 1990; Stenström et al., 2002; Columbus, 2010; Achiri-Taboh, 2015; Criado Peña, 2016; among others). In addition, Table 6 shows that there is an overall direct proportional relation between frequency of functional type and marker variation. It can be observed that informational BEITQs display eleven of the thirteen BEITQ types detected (with the exception of *isn't that right?* and *see?*), and the attitudinal category nine (with no records of *am I right*, *don't you think*, *is that right?* and *or what?*). In contrast, the other functional types register a narrower range of BEITQs: hortatory (*ok?*, *eh?*, *right?*, *yeah?*), focusing (*right?*, *you see?*, *all right?*), phatic (*ok?*, *right?*) and regulatory (*is that right?*). All in all, the most frequent BEITQ is *ok?* (26%), followed by *right?* (19.7%), and *yeah?* (12.0%). The remaining ten types register percentages below 10%.

In the case of EPITQs, Table 7 displays the seventeen types registered in the corpus out of the twenty-five. No records were found for (*estás de*) *acordo?*, *está certo?*, *eh?* (with 10 BEITQ tokens), *não é assim?*, *né?*, *não está certo?*, *sim?*, *verdade?*, *vês?*. The fact that no instance of *isn't that so?* or *não é assim?* were found suggests a more limited use in the two languages. Strikingly, unlike what happened in the English data set, 100% of the occurrences of the regulatory use (162), the prevalent one in (European) Portuguese, were performed by just one IQT, i.e., *não é?*.

This result provides evidence of the grammaticalization of *não é?* as a textual device, as well as of its pervasive use. It registers scores for six of the seven functionalities under inspection, a tendency already perceived by Cruz-Ferreira (1981). The two other textual uses, phatic and focusing, display more variation with eight (*está a ver?*, *ok?*, *certo?*, *correcto?*, *está bem?*, *hã?*, *não é?*, *não é verdade?*) and three IQTs types (*não é?*, *hã?*, *está a ver?*), respectively. Turning

Table 6
Functional types of BETQIQTs.

BEIQT N = 142	attitudinal N = 45	focusing N = 12	hortatory N = 5	information N = 75	phatic N = 4	Regulatory N = 1
ok? 37 (26%)	25 (55.6%)	(-)	2 (40.0%)	7 (9.3%)	3 (75.0%)	(-)
right? 28/20%	4 (8.9%)	5 (41.7%)	1 (20.0%)	17 (22.7%)	1 (25.0%)	(-)
yeah? 17/12%	2 (4.4%)	(-)	1 (20.0%)	14 (18.7%)	(-)	(-)
you see? 14/9.9%	6 (13.3%)	5 (41.7%)	(-)	3 (4.0%)	(-)	(-)
eh? 10/7.0%	2 (4.4%)	(-)	1 (20.0%)	7 (9.3%)	(-)	(-)
is that right? 9/6.3%	(-)	(-)	(-)	8 (10.7%)	(-)	1 (100.0%)
or not? 9/6.3%	1 (2.2%)	(-)	(-)	8 (10.7%)	(-)	(-)
all right? 6/4.2%	2 (4.4%)	2 (16.7%)	(-)	2 (2.7%)	(-)	(-)
don't you think? 4/2.8%	(-)	(-)	(-)	4 (5.3%)	(-)	(-)
or what? 4/2.8%	(-)	(-)	(-)	4 (5.3%)	(-)	(-)
see? 2/1.4%	2 (4.4%)	(-)	(-)	(-)	(-)	(-)
am I right? 1/0.7%	(-)	(-)	(-)	1 (1.3%)	(-)	(-)
isn't that right? 1/0.7%	1 (2.2%)	(-)	(-)	(-)	(-)	(-)

Table 7
Functional types of EPIQITs.

EPIQIT	attitud (N = 67)	challeng (N = 1)	facilitat (N = 3)	focusing (N = 11)	hortatory (N = 11)	informat (N = 91)	phatic (N = 51)	regulat (N = 162)
está a ver?	(-)	(-)	(-)	1 (9.1%)	(-)	(-)	16 (31.4%)	(-)
não acha?	2 (3.0%)	(-)	2 (66.7%)	(-)	(-)	(-)	(-)	(-)
ok?	(-)	(-)	(-)	(-)	(-)	(-)	16 (31.4%)	(-)
certo?	2 (3.0%)	(-)	(-)	(-)	(-)	1 (1.1%)	11 (21.6%)	(-)
correcto?	(-)	(-)	(-)	(-)	(-)	(-)	1 (2.0%)	(-)
está bem?	(-)	(-)	(-)	(-)	7 (63.6%)	3 (3.3%)	1 (2.0%)	(-)
está okay?	(-)	(-)	(-)	(-)	1 (9.1%)	(-)	(-)	(-)
hã?	(-)	(-)	(-)	2 (18.2%)	(-)	(-)	1 (2.0%)	(-)
hhh?	(-)	(-)	(-)	(-)	(-)	1 (1.1%)	(-)	(-)
não?	1 (1.5%)	(-)	(-)	(-)	(-)	28 (30.8%)	(-)	(-)
não é?	56 (83.6%)	1 (100.0%)	(-)	8 (72.7%)	2 (18.2%)	33 (36.3%)	3 (5.9%)	162 (100.0%)
não é verdade?	3 (4.5%)	(-)	(-)	(-)	1 (9.1%)	(-)	2 (3.9%)	(-)
não já?	(-)	(-)	(-)	(-)	(-)	1 (1.1%)	(-)	(-)
não te parece?	(-)	(-)	1 (33.3%)	(-)	(-)	(-)	(-)	(-)
ou não?	(-)	(-)	(-)	(-)	(-)	15 (16.5%)	(-)	(-)
ou quê?	3 (4.5%)	(-)	(-)	(-)	(-)	1 (1.1%)	(-)	(-)
pois não?	(-)	(-)	(-)	(-)	(-)	8 (8.8%)	(-)	(-)

Table 8

The functions of equivalent EPIQT and EPIQT.

BEIQTs	EPIQTs	function
ok?	okay?	≠
right?	certo?/ correto?	≠
you see?	está(s) a ver?	≠
eh?	hã?	≠
	hhh?	informational
or not?	ou não?	informational
all right?	está bem?	≠
don't you think?	não acha(s)?/ não te parece?	≠
or what?	ou quê?	≠
isn't that right?	não é verdade?	attitudinal

to the second most frequent functionality, the informational type also exhibits marker variation, but is mainly expressed by negative ITQs (92.4%), i.e., *não é?*, with the highest score, *não?*, *ou não?* and *pois não?*. The remaining 7.6% corresponds to two positive (*está bem?*, *certo?*) and one disjunctive (*ou quê?*) informational IQT. Hortatory (*está bem*, *está okay?*, *não?*, *não é verdade?*) and challenging (*não é?*) EPIQTs also confirm a propensity towards negative polarity IQT. This tendency was also observed in the variable type in this language (Gómez González, 2014), which confirms the claim that in Portuguese questions tags tend to express negative polarity, explaining why they are generally referred as *marcadores de negação frásica* 'sentential negative markers' (Gonçalves, 1995).

Lastly, the data also reveal that, while some IQTs seem to have an inherent function (*am I right?*, *don't you think?*, *or what?* – informational; *see?* – attitudinal; *pois não?*, *ou não?* – informational); others are more polyfunctional such as *ok?* (attitudinal, informational, phatic, hortatory), *you see?* (informational, affective, focusing), *não é?* (registering seven, out of eight, functionalities) or *não é verdade?* (affective, phatic and hortatory). However, it should be observed that, out of the nine pairs of equivalent IQTs detected in our sample, only three of them share the same most frequent function, as shown in Table 8 below. This seems to suggest that there are no one-to-one form-to-function correspondences between IQT types across the two languages.

4.2. RQ # 2: Grammatical characteristics of functional types

Tables 9 and 10 (Appendix B) present the quantitative information corresponding to the grammatical characteristics of the proposed functional types. Comparing the results, firstly, the numbers suggest that both BEIQTs and EPIQTs, but especially the former, are normally appended to, rather than inserted within, their ANC. The only exception is the focusing type that registers the highest incidence of medial occurrences. This feature profiles a contrastive characteristic in the tag question paradigm as the same positional tendencies have been observed for variable tag questions across the two languages, even if IQTs show a more skewed distribution (Gómez González, 2014, 2018).

Secondly, considering (meta)polarity, the data uncover a stark contrast between the two data sets. In BEITQs, constant positive polarity values are the most common across the six functional types (76.8% BE vs. 26.2% EP), no records being registered of the constant negative pattern. Specifically, constant positive polarity schemes occur in 100% of the occurrences of focusing and regulatory BEITQs, as well as in 80% of the hortatory type. The remaining patterns, including all the occurrences of reversed patterns, cluster around the most frequent uses, i.e., information and attitudinal, in keeping with the wider variation they also exhibit in other parameters of analysis. In EPITQs, on the other hand, reversed polarity constructions are the general trend (72.0% EP vs. 13.4% BE), predominantly (+ / -) patterns. They represent 70.7% of the Portuguese data set, as well as the majority of attitudinal (73.1%), facilitative (100%), focusing (63.6%) informational (79.1%) and regulatory (87.7%) uses. Alternatively, constant (+ / +) patterns (13.6%) are only found in three uses, most of them belonging to the phatic (82.4%) or hortatory (63.6%) types and only one token of the focusing type. Constant (- / -) schemes (12.3%) are registered in more uses, but with lower frequencies overall: attitudinal (19.4%), informational (14.3%), regulatory (12.3%), focusing (9.1%) and only one token in the focusing, phatic and challenging (the only one) categories. Regarding neutral polarity, the ITQs seem to confirm the same tendency across functional types in the two languages: there are more instances of positive neutral (7.7% BE vs. 1.8% EP) than negative neutral (3% BE vs. 2.1% EP).

Zooming out these (meta)polarity results, three main contrastive conclusions are to be drawn in the light of relevant literature. One is that in English invariable and variable question tags seem to deploy two complementary strategies

concerning polarity: the former generally keep it constant, while the latter mostly reverse it (Aijmer, 1979; Biber et al., 1999: 208; Quirk et al., 1985; Tottie and Hoffmann, 2006, 2009; Axelsson, 2011b; Gómez González, 2012, 2014, 2016, 2018; Kimps et al., 2014a; Kimps, 2018; Barron et al., 2015). The second conclusion is that this tendency cannot be extrapolated to languages like Portuguese (nor possibly to other Romance languages), where negative QTs, whether variable or invariable, are the general rule as already noted in the previous section. And thirdly, confirmations and deviations from these tendencies have a functional underpinning (Biber and Egbert, 2016; Gómez González, 2014, 2018). Broadly, in reversed patterns negative question tags are used to assess the anchor proposition as *unexpected* (DeLancey, 2001: 370; Simon-Vandenberg and Aijmer, 2007; Kimps, 2018) or as *self-evident* (Brazil, 2008: 36) or as an alternative negative option in disjunctive tag questions, typically in informational and attitudinal uses in English and as regulatory strategies in Portuguese. In contrast, polarity-neutral and affirmative ITQs are used, particularly in English, as reinforcement strategies of a previous, usually but not necessarily, affirmative statement; or, otherwise, they are generally associated with speakers' negotiations of desired actions, as illustrated in (13), where four IQTs are stacked (*yeah OK yeah right*) to involve the addressee in the proposal, to which s/he agrees.

Moving on to the fourth parameter, Mood, the large majority of ITQs in the two data sets, but particularly in Portuguese, have declarative anchors. This is the only mood pattern that deploys all the functional capabilities recognized in each language. Interrogative anchors, on the other hand, are much less frequent and are mostly found in informational ITQs, and less frequently in attitudinal ITQs in the two data sets. Even less frequent are imperative anchors. They are restricted to hortatory EPITQs and also represent 100% of such use in BEITQs. In the English sample, however, imperative anchors are also found to deploy attitudinal and phatic strategies. Exclamative anchors only occur in one BEITQ with *eh?* – (4e).

To bring this section to an end, some regularities should be observed with regard to the remaining grammatical characteristics of ITQs across the two languages. Most anchors are not truncated, especially in Portuguese (90% BEITQs and 96% EPITQs). In English, only the anchor of informational tag questions is truncated, but they again represent a minority in this use (18.7%). In Portuguese, truncated anchors generally display percentages below 14%, but across a wider array of functions (facilitative, focusing, hortatory, informational, phatic and regulatory). Likewise, most anchors are realized by simple clauses with present (particularly in English) or past (especially in Portuguese) tenses. However, complex-subordinated sentences prevail in BEITQs with an attitudinal, focusing, and phatic function, as well as in the focusing, hortatory and phatic EPITQs. As already noted, it is the most frequent uses that display a wider tense selection in the two samples (informational and attitudinal in BE vs. regulatory and informational in EP), while choices are more constrained in less frequent uses (focusing, regulatory, phatic in BE vs. facilitative and challenging in EP), hortatory ITQs being associated with imperative ANCs. Lastly, the IQT Subject is reported as irrelevant (null realizations) or not applicable in most cases across the two languages.

The above results partly confirm those obtained in previous investigations concerning VTQs in English (Axelsson, 2011b; Gómez González, 2012, 2016; Tottie and Hoffmann, 2006; Kimps, 2018, *inter alios*), contrasting VTQs across languages (Gómez González, 2014) or comparing variable and invariable tag questions in English (Gómez González, 2018). The fact that most variable and invariable TQs in English and Portuguese have a declarative anchor, typical of statements, followed by an interrogative pattern in the QT, associated with questions, underscores the unmarked blending of formal features and speech. At the semantic-pragmatic level, this blending suggests that speakers first use declarative anchors to encode meaning from a knowledgeable status, and then use (in)variable question tags to express some type of meaning imbalance that, at least formally, expects a response in order to level out such an imbalance. With this strategy, at least formally, speakers present themselves as *secondary knowers* or *secondary actors*, and turn to the co-participants as *primary knowers* or *primary actors*, sources or beneficiaries/addressees of the goals pursued in the communicative acts (Berry, 1981a, 1981b, 2016; Martin, 1992; Heritage et al., 2011; Kimps, 2018; Gómez González, 2018, 2020).

Nevertheless, the findings partly contravene the assumption that unmarked variable and invariable TQs share similar characteristics and are always interchangeable. In addition to the functional disparities already noted, there also exist grammatical divergencies that distinguish invariable and variable tag questions, for instance, those concerning such parameters as ANC Tense, Type of Subject (ANC), Type of Subject (IQT) and IQT Tense. They are mostly related to the formal make-up of the constructions and to the constraints imposed by the features of each language.

4.3. RQ # 3: Dialogic and generic features

Considering the dialogic level, Tables 11 and 12 show that, although all in all ITQs tend to receive no verbal response both in the English (73.2%) and Portuguese (66.3%) data sets, other/self replies are registered particularly in the informational and attitudinal uses, a tendency that is more marked in Portuguese.

Table 11

Dialogic / Generic features of functional types of BEIQTs.

	attitudinal (N = 45)	focusing (N = 12)	hortatory (N = 5)	informational (N = 75)	phatic (N = 4)	regulatory (N = 1)	Total (N = 142)
Response							
no	42 (93.3%)	11 (91.7%)	4 (80.0%)	42 (56.0%)	4 (100.0%)	1 (100.0%)	104 (73.2%)
other reply	3 (6.7%)	(–)	1 (20.0%)	30 (40.0%)	(–)	(–)	34 (23.9%)
self reply	(–)	1 (8.3%)	(–)	3 (4.0%)	(–)	(–)	4 (2.9%)
Genre							
private	17 (37.8%)	7 (58.3%)	4 (80.0%)	50 (66.7%)	(–)	1 (100.0%)	79 (55.6%)
dialog	0.8 nfpttw	0.3 nfpttw	0.2 nfpttw	2.4 nfpttw		0.04 nfpttw	3.8 nfpttw
205,608 w							
public	28 (62.2%)	5 (41.7%)	1 (20.0%)	25 (33.3%)	4 (100.0%)	(–)	63 (44.4%)
dialog	1.6 nfpttw	0.3 nfpttw	0.05 nfpttw	1.4 nfpttw	0.2 nfpttw		3.6 nfpttw
171,059 w							
Monologic	(–)	(–)	(–)	(–)	(–)	(–)	(–)
260,895 w							

Table 12

Dialogic / Generic features of functional types of EPITQs.

	attitudinal (N = 67)	challeng (N = 1)	facilit (N = 3)	focusing (N = 11)	hortat (N = 11)	informat (N = 91)	phatic (N = 51)	regulat (N = 162)	Total (N = 397)
Response									
no	49 (73.1%)	1 (100.0%)	(–)	10 (90.9%)	4 (36.4%)	9 (9.9%)	39 (76.5%)	151 (93.2%)	263 (66.2%)
other reply	15 (22.4%)	(–)	3 (100.0%)	1 (9.1%)	7 (63.6%)	81 (89%)	12 (23.5%)	9 (5.6%)	128 (32.2%)
self reply	3 (4.5%)	(–)	(–)	(–)	(–)	1 (1.1%)	(–)	2 (1.2%)	6 (1.6%)
Genre									
private	35 (52.2%)	1 (100.0%)	2 (66.7%)	1 (9.1%)	7 (63.6%)	54 (59.3%)	7 (13.7%)	41 (25.3%)	148 (37.3%)
dialogic	3.1 nfpttw	0.1 nfpttw	0.2 nfpttw	0.1 nfpttw	0.6 nfpttw	4.8 nfpttw	0.6 nfpttw	3.7 nfpttw	13.3 nfpttw
111,264 w									
public dialogic	16 (23.9%)	(–)	1 (33.3%)	6 (54.5%)	4 (36.4%)	28 (30.8%)	33 (64.7%)	51 (31.5%)	139 (35%)
107,818 w	1.5 nfpttw		0.1 nfpttw	0.5 nfpttw	0.4 nfpttw	2.6 nfpttw	3.1 nfpttw	4.7 nfpttw	12.9 nfpttw
monologic	16 (23.9%)	(–)	(–)	4 (36.4%)	(–)	9 (9.9%)	11 (21.6%)	70 (43.2%)	110 (27.7%)
165,041 w	1 nfpttw			0.2 nfpttw		0.5 nfpttw	0.7 nfpttw	4.2 nfpttw	6.6 nfpttw

This finding indicates that invariable and variable tag questions behave somewhat differently across the two languages. According to previous investigations (Gómez González, 2014, 2018), the latter show a slightly higher tendency to obtain verbal responses. The lack of verbal responses in ITQs confirms the bleaching of their questioning function in usage, especially in interactive or textual uses across the two languages. However, although a minority (26.8% BEIQT and 33.8% EPITQs), there are also tokens evidencing the dialogic and conducive nature of interactional (information-seeking or facilitative) invariable tag questions. These generally *project* a next turn as the second-pair part of an adjacency pair, so that a confirmation is provided – as in (13c), (15), (20a) –, but less frequently such dispreferred seconds as refutations – as in (20b) – may be offered (Ford et al., 2003; Axelsson, 2014). Further illustrations of the rhetorical nature of, especially interactive, ITQs are presented in (21), all self-replied not allowing the addressee to intervene.

-
- (21) a. A: Debilitating **right**? Yes. Uhm. Uhm. < ICE-GB:S1A-061 #080:1:A>
 b. NAT: há quarenta e tal anos / **não é ?** pois / há quarentas e & mui / quarenta e cinco para aí
 <CORALRpfamnn15>
 NAT: there-are forty and so years not is-PRESENT-3S? right / there-are forty and & many / forty five more or less
 NAT: 'forty or so years ago, isn't it? Right, forty and many, forty five more or less'
-

Turning to the contextual level, the scrutiny of [Tables 11 and 12](#) confirms the dialogic nature of ITQs since they are far more used in dialogic than in monologic texts in the two languages. However, the frequency rate of the constructions in monologues is much higher in the Portuguese data set (27%, 6.6 nfptw), particularly in the regulatory use (43.2%, 4.2 nfptw), because no record was found in the English sample. Instead, BEITQs were evenly distributed between private (N = 79, 55.6%, 3.8 nfptw) and public (N = 63, 44.4%, 3.6 nfptw) dialogues. EPITQs, in contrast, exhibited a more balanced distribution across the three genres: private dialogues also registered the highest percentage (37.3%, 13.3 nfptw), followed by public dialogic discourse (35%, 12.9 nfptw) and monologues (27.7%, 6.6 nfptw). On closer inspection, it can also be observed that, in BEITQs, only the attitudinal type registers a higher percentage in public than in private dialogues, while in the other functionalities the tendency is reversed. Turning to the Portuguese sample, more common in private dialogic discourse are those serving interactional purposes (attitudinal, challenging, facilitative hortatory and informational), while interactive EPITQs (focusing, phatic and regulatory) register higher percentages in public dialogic discourse.

These results uncover yet another finding of a comparative nature that is relevant to describe the tag question paradigm. In prior work ([Gómez González, 2014: 21](#)), variable tag questions were reported to exhibit a similar generic distribution in the two languages under inspection. It appears then that in BE both variable and invariable tag questions tend to be used in private informal conversations, whereas in EP tag questions seem to be more frequent in all domains, including monologues. But, while according to [Gómez González \(2014\)](#) variable tag questions in Portuguese tend to be more pervasive in public contexts, our data reveal that EPITQs show slightly higher scores in private dialogues.

More generally, the data seem to support the idea that it is the context and the participants' roles that explain the choice of a particular type and function of TQ, as also suggested in some prior investigations (e.g., [Calnan and Davidson, 1998: 31](#); [Cameron et al., 1988: 86](#); [Cheng and Warren, 2001: 1436](#); [Gómez González, 2012, 2014, 2016, 2018](#); [Kimps, 2018](#)). A more detailed study of this issue, however, goes beyond the scope of this contribution.

4.4. RQ # 4: Sociolinguistic variation

[Tables 13 and 14](#) summarize our data for the ITQ user's Sex and Age parameters. Regarding Sex, 64.8% BEITQs and 57.4% EPITQs were used by men.⁹ This result is in keeping with the expected distribution as in both corpora male discourse predominates. In the spoken component of ICE-GB, 460,967 words are uttered by male speakers, while the female set registers only 176,595 words. In C-ORAL-ROM, the proportion is 144 male and 117 female speakers. However, if we consider the nfptw values in ICE-GB, it turns out that ITQs are more frequent overall in female discourse by 0.8 points (except for the phatic, regulatory and focusing types), while in the Portuguese data set women users outnumber males in three functionalities (attitudinal, challenging and hortatory). This is in keeping with some prior studies ([Lakoff, 1975](#); [Holmes, 1984](#); [Cameron et al., 1988](#); [Eckert and McConnell-Ginet, 2003](#)), but it contravenes the conclusions of other investigations (e.g., [Dubois and Crouch, 1975](#); [Lapadat and Seesahai, 1977](#)).

Finally, considering the age of ITQ users, [Tables 13 and 14](#) seem to profile an age trend in the two languages because the bulk of BEITQs and EPITQs cluster around the 46–65 and 41–60 age groups across most functional types, which are also the age groups sampled with more speakers. Nevertheless, if the nfptw values in ICE-GB are observed, it turns out that speakers aged 18–25 use ITQs twice or more as much than the other age groups. In EPITQs, deviations are detected in the challenging and phatic functional types, which show more frequent uses by the 26–40 age rank.

⁹ Neither ICE-GB nor C-ORAL-ROM provide a way to obtain subcorpora according to sociolinguistic variables only, nor is there published information available on the corpus compilation statistics concerning metadata for each and all speakers. Furthermore, a number of ITQs had to be assigned to the "Unknown" Age category across the two corpora (BEITQ N = 10 and EPEITQ N = 40) because the user's age was not offered in those cases. Nevertheless, in C-ORAL-ROM we could obtain the number of speakers per sex and age (but not the number of words) for each set, while in ICE-GB we could calculate nfptw in each category based on gender and age variable searches in ICECUP 3.1 combined with WordCounts for each file after removing coding material. This is not an ideal situation in terms of this study, but it helps overcome the issues derived from the imperfect information in the corpora, which are ultimately in the hands of the corpora compilers.

Table 13

Sociolinguistic variation across BEITQ functional types.

	attitudinal (N = 45) nfpttw	focusing (N = 12) nfpttw	hortatory (N = 5) nfpttw	informational (N = 75) nfpttw	phatic (N = 4) nfpttw	regulatory (N = 1) nfpttw	Total (N =) nfpttw
Sex							
male	30 (66.7%)	10 (83.3%)	3 (60.0%)	44 (58.7%)	4 (100.0%)	1 (100.0%)	92 (64.8%)
460,967 w	0.6	0.2	0.06	0.9	0.08	0.02	2
female	15 (33.3%)	2 (16.7%)	2 (40.0%)	31 (41.3%)	(–)	(–)	50 (35.2%)
176,595 w	0.8	0.1	0.1	1.7			2.8
Age							
18–25	6 (13.3%)	1 (8.3%)	3 (60.0%)	23 (30.7%)	(–)	1 (100.0%)	34 (23.9%)
82,731 w	0.7	0.1	0.3	2.8		0.1	4.1
26–45	17 (37.8%)	2 (16.7%)	(–)	27 (36.0%)	(–)	(–)	46 (32.4%)
316,506 w	0.5	0.06		0.8			1.4
46–65	17 (37.8%)	7 (58.3%)	1 (20.0%)	18 (24.0%)	4 (100.0%)	(–)	47 (33.1%)
219,085 w	0.8	0.3	0.04	0.8	0.2		2.1
over 66	4 (8.9%)	(–)	1 (20.0%)	(–)	(–)	(–)	5 (3.5%)
19,240 w	2.07		0.5				2.6
Unknown	1 (2.2%)	2 (16.7%)	(–)	7 (9.3%)	(–)	(–)	10 (7.1%)

Table 14

Sociolinguistic variation across EPITQ functional types.

	attitudinal (N = 67)	challeng (N = 1)	facilit (N = 3)	focusing (N = 11)	hortat (N = 11)	informat (N = 91)	phatic (N = 51)	regulat (N = 162)	Total
Sex									
male	33 (49.3%)	(–)	3 (100.0%)	9 (81.8%)	4 (36.4%)	53 (58.2%)	42 (82.4%)	84 (51.9%)	228 (57.4%)
N = 144									
female	34 (50.7%)	1 (100.0%)	(–)	2 (18.2%)	7 (63.6%)	38 (41.8%)	9 (17.6%)	78 (48.1%)	169 (42.6%)
N = 117									
Age									
18–25	7 (10.4%)	(–)	(–)	1 (9.1%)	(–)	17 (18.7%)	1 (2.0%)	8 (4.9%)	34 (8.5%)
N = 37									
26–40	21 (31.3%)	1 (100.0%)	(–)	2 (18.2%)	2 (18.2%)	16 (17.6%)	26 (51.0%)	30 (18.5%)	98 (24.7%)
N = 40									
41–60	29 (43.3%)	(–)	2 (66.7%)	2 (18.2%)	8 (72.7%)	37 (40.7%)	19 (37.3%)	88 (54.3%)	185 (46.6%)
N = 69									
over 60	8 (11.9%)	(–)	1 (33.3%)	3 (27.3%)	(–)	12 (13.2%)	1 (2.0%)	15 (9.3%)	40 (10.1%)
N = 8									
Unknown	2 (3.0%)	(–)	(–)	3 (27.3%)	1 (9.1%)	9 (9.9%)	4 (7.8%)	21 (13.0%)	40 (10.1%)
N = 107									

Interestingly, despite the fact speakers whose age is *unknown* are the most numerous, they register the lowest relative incidence of EPITQ use.

These findings, in particular those reported for BEITQ, seem to confirm the conclusions of prior investigations noting that ITQs tend to be used by teenagers and younger speakers (Stenström, 1997, 2005; Stenström et al., 2002; Andersen, 2001; Tottie and Hoffmann, 2006; Moore and Podesva, 2009; Kimps et al., 2014b; Kimps, 2018). Nevertheless, these corpus-based differences seem to be too small to attribute the age variable a significant impact on the choice of ITQs functional types, as also concluded in other studies (Kimps et al., 2014b; Kimps, 2018; Gómez González, 2018), among others.

4.5. RQ # 5: Significance of results and model of predicted associations

As can be seen in Table 15, all fifteen variables have a significant association with Function in the Portuguese data set, according to both *p*-values and adjusted *p*-values. In English, in contrast, *p*-values register significant associations

Table 15
Significance of statistical analyses (p-value, CC and tau for *Function*).

Variables	BEITQs							EPITQs						
	p-value	Signif	Adjusted p-value	Signif	CC	tauxy	tauyx	p-value	Signif	Adjusted p-value	Signif	CC	tauxy	tauyx
IQT type	0.031	*	0.0759	ns	0.658	0.116	0.298	0.000	***	0.0008	***	0.852	0.361	0.408
IQT position	0.002	**	0.0150	*	0.489	0.314	0.058	0.000	***	0.0008	***	0.612	0.599	0.028
polarity	0.482	ns	0.4818	ns	0.336	0.071	0.063	0.003	**	0.0037	**	0.672	0.330	0.162
meta-polarity	0.105	ns	0.1431	ns	0.327	0.081	0.060	0.000	***	0.0008	***	0.514	0.289	0.074
Type of Subject (ANC)	0.357	ns	0.3828	ns	0.354	0.036	0.072	0.029	*	0.0290	*	0.387	0.024	0.029
Type of Subject (IQT)	0.183	ns	0.2286	ns	0.304	0.066	0.024	0.000	***	0.0008	***	0.462	0.271	0.048
Mood_ (ANC)	0.006	**	0.0244	*	0.665	0.287	0.097	0.000	***	0.0008	***	0.632	0.290	0.042
IQT tense	0.097	ns	0.1431	ns	0.244	0.063	0.004	0.000	***	0.0008	***	0.650	0.470	0.211
ANC tense	0.006	**	0.0244	*	0.731	0.114	0.171	0.004	**	0.0046	**	0.676	0.132	0.127
Truncated ANC	0.040	*	0.0759	ns	0.298	0.098	0.059	0.026	*	0.0284	*	0.265	0.076	0.014
Response	0.039	*	0.0759	ns	0.403	0.160	0.109	0.003	**	0.0037	**	0.640	0.394	0.203
Clause type (ANC)	0.050	*	0.0833	ns	0.350	0.103	0.044	0.000	***	0.0008	***	0.386	0.043	0.030
Sex	0.337	ns	0.3828	ns	0.199	0.041	0.010	0.000	***	0.0008	***	0.244	0.063	0.010
Age	0.037	*	0.0759	ns	0.475	0.081	0.083	0.000	***	0.0008	***	0.383	0.047	0.037
Genre	0.002	**	0.0150	*	0.323	0.117	0.046	0.000	***	0.0008	***	0.417	0.106	0.061

p-values:***: p-value ≤ 0.001 , **: $0.001 < \text{p-value} \leq 0.01$, *: $0.01 < \text{p-value} \leq 0.05$, ns: p-value > 0.05 .

Table 16
Feature importance for BEITQs and EPITQs.

Rank	BEITQs	EPITQs
1	IQT type	Response
2	Response	IQT type
3	Mood (ANC)	IQT tense
4	IQT position	IQT position
5	Genres	Polarity
6	Truncated (ANC)	Mood(anchor)
7	Type of subject (ANC)	Genres
8	ANC tense	ANC tense
9	Type of subject (IQT)	Type of subject (IQT)
10	IQT tense	Age
11	Polarity	Type of subject (ANC)
12	Age	Sex
13	Clause type (ANC)	Clause type (ANC)
14	Sex	Truncated (ANC)

only in nine variables, i.e., IQT position, Mood (ANC), ANC tense, Genre, IQT type, Truncated ANC, Response, Clause type (ANC) and Age, although only the first four remain significant if adjusted *p*-values are considered.

Furthermore, if we consider Table 16 ranking the CCs showing significant association, the variables that register higher values (both in BEITQs and EPITQs) are IQT type, Response, Mood (ANC), IQT position, and Anchor tense.¹⁰ Considering Goodman and Kruskal tau's, although values are very low in the English data set (with the exception of tag

¹⁰ The variable Meta-polarity is eliminated from the model since it is constructed upon Polarity.

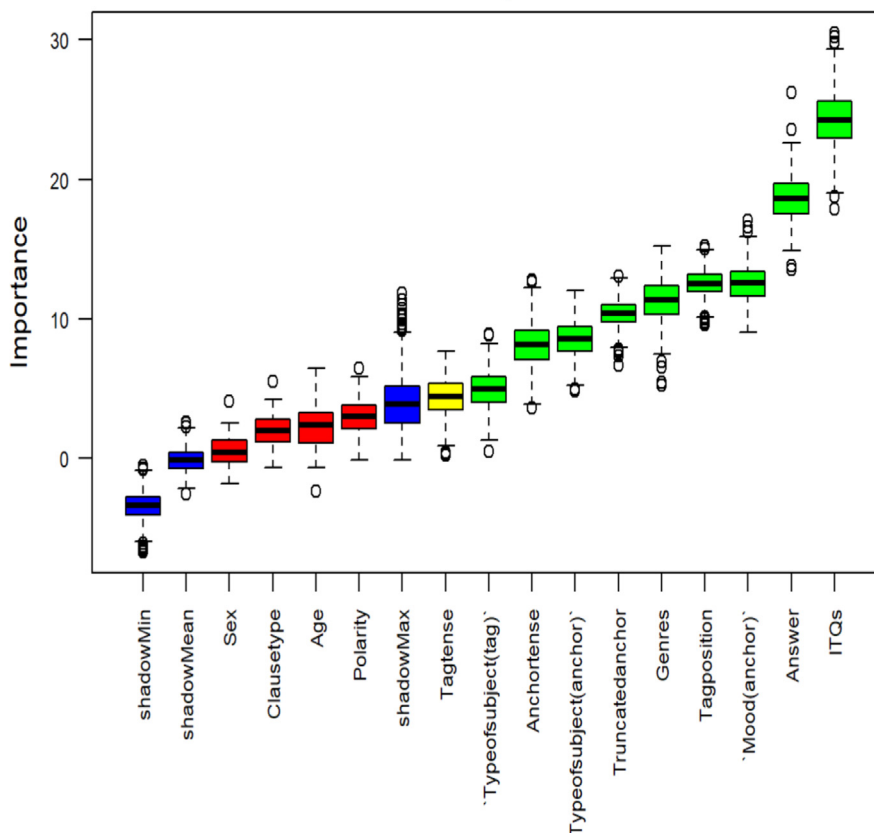


Fig. 2. Feature importance for BEITQs.

position), forward associations (from Function to variables) display higher values and therefore seem to generally prevail, particularly in EPITQs.

Additionally, Figs. 2 and 3 display the boxplots featuring the models of predicted associations for the characteristics of EPITQs and BEITQs found in our sample based on the Boruta algorithm with 500 iterations. The colours green, yellow and red represent scores of statistically confirmed, tentative and rejected features, respectively.¹¹ Accordingly, Fig. 2 shows that the values of the nine features highlighted in green can be predicted for the functional types of BEITQs, whereas Age, Polarity, Tag tense, Clause type and Sex are rejected. Turning to EPITQs, Fig. 3 indicates that the values of eleven out of the fourteen variables under study (except for Clause type, Sex and Truncated ANC) can be forecast for the proposed functional types.

To conclude, the most important variables in both data sets are Response and ITQ type, though in reverse order. Importance values are higher in the Portuguese data set, in general, where forward associations (from Function to variables) are higher than backward ones (from variables to Function).

¹¹ Additionally, blue represents a shadow feature. A shadow feature is a feature created by randomly permuting each variable in the data frame, and they are used to set the score threshold for feature importance, allowing to confirm or reject the original variables. A variable is of predictive usefulness when its Z score (i.e., the equivalent of importance score for shadow features) is higher than the maximum Z score of the shadow feature. In the boxplots displayed in Figs. 2 and 3, the terms *shadow mean*, *shadow min* and *shadow max* represent the distribution of the mean, minimum and maximum values of z score for shadow features over the 500 runs of the Boruta algorithm, respectively.

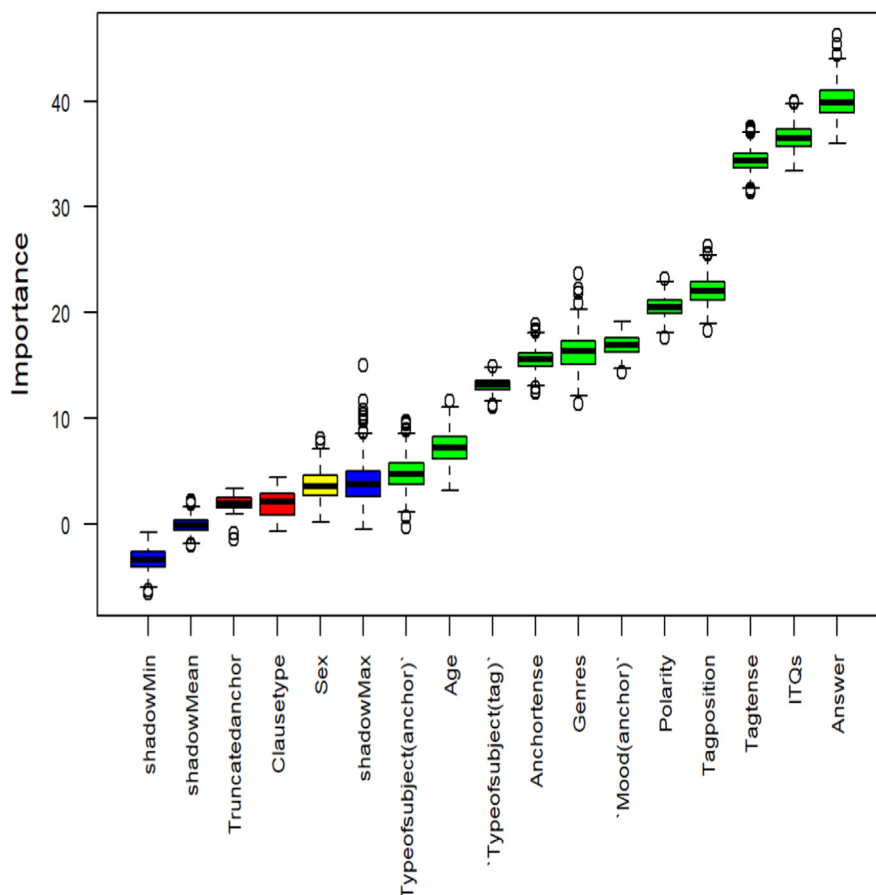


Fig. 3. Feature importance for EPITQs.

5. CONCLUSIONS

This study has offered a comparative characterization of invariable tag questions in British English and European Portuguese. We have claimed that, although invariable and variable question tags can be modelled within the same functional paradigm, they may not always be interchangeable in the same contexts mostly due to functional motivations, although grammatical disparities have also been observed. In addition, qualitative and quantitative corpus-driven data have been examined on eight functional types of invariable tag questions according to fifteen variables (Appendix A): i.e., *informational*, *affective*, *challenging*, *hortatory*, *facilitative*, *focusing*, *phatic* and *regulatory*.

Using statistical significance tests, it has been determined that all the observed variables have a significant association with the proposed functional types of invariable tag questions in Portuguese, whereas in English only nine categories are significant: Tag position, Mood (ANC), ANC tense, Genre, IQT type, Truncated anchor, Response, Clause type (ANC) and Age. However, only the first four remain significant in the two languages if adjusted *p*-values are considered. Furthermore, comparing Contingency Coefficients, the highest values in both data sets are Response, ITQ type, Tag position, ANC tense and Mood (ANC), the first two being the most important variables in the two languages (though in reverse order). Goodman and Kruskal's tau measures are higher in Portuguese, while they are very low in English with the exception of IQT position. In both languages, particularly in Portuguese, forward associations ($\tau_{xy}(x, y)$) prevail, which suggests that Function determines the features of ITQs across the independent variables, excluding Clause type (ANC), Sex and Truncated (ANC) that were rejected.

Based on these results and on prior work, four overarching conclusions can be drawn. The first is that tag questions, whether invariant or variant, are not only more frequently used, but also more functionally varied in Portuguese than in English. Secondly, while in Portuguese both invariant or variant tag question are predominantly used as interactive or textual strategies, in English the two types are mostly used in their interactional capacities (epistemic and attitudinal).

Thirdly, we have seen that, while in Portuguese negative tag questions generally prevail (as well as in other Romance languages), in English invariable and variable question tags seem to deploy two complementary polarity strategies, the former generally keeping it constant (positive) and the latter mostly reversing it. Additionally, evidence has been provided that there are no one-to-one form-to-function correspondences between the nine equivalent invariable tag questions pairs detected in our English and Portuguese data sets.

Finally, the Boruta algorithm has been implemented to propose corresponding models of predicted associations with nine and eleven variables for invariable tag questions in English and Portuguese, respectively. More specifically, we have explained that the functions performed by invariable tag questions, especially in Portuguese, determine whether a response is provided or not, as well as the ITQ type, the position of the IQT, alongside the Tense and Mood characteristics of the ANC. All the other variables display a weaker or no significant association with the functional types of invariable tag questions investigated.

Given the richness of the data and the theoretical breadth of the tag question paradigm, further research should continue to explore the pragmatic peculiarities of each and every type of invariable tag question, either for their own sake or in contrast with the variable type or equivalent constructions across varieties and languages. Similarly, a more thorough analysis focusing on speaker’s variables (e.g., role in interaction, gender and age) should be conducted in order to be able to establish on firmer grounds the correlation of such variables to invariable tag question use.

ACKNOWLEDGEMENTS

The authors wish to gratefully thank Chris Butler, Karin Aijmer, Karin Axelsson, Nicole Dehé and Lachlan Mackenzie for fruitful discussions, as well as the USC Statistical Consulting Service for useful input on this work. The research was funded by the Spanish National Research Agency (FFI2013-40517-P; 2019-PN145) and the Xunta de Galicia (GRC2015/002 GI-1924; ED431C 2019/01), whose financial support is hereby acknowledged. The usual disclaimers apply.

APPENDIX A

Variables and values used in the empirical analysis

A. Dependent variable
1. Function
1 = affective
2 = challenging
3 = facilitative
4 = focusing
5 = hortatory
6 = informational
7 = phatic
8 = regulatory
C. Independent variables
Bi. Grammatical
2. IQT type (cf. Table 2)
3. IQT position
1 = final
2 = medial
4. Meta-polarity
1 = reversed (1 = positive–negative, 2 = negative–positive)
2 = constant (3 = positive–positive, 4 = negative–negative)
3 = neutral (5 = positive–neutral, 6 = negative–neutral)

5. Polarity

- 1 = positive-negative
- 2 = negative-positive
- 3 = positive-positive
- 4 = negative-negative
- 5 = positive-neutral
- 6 = negative-neutral

6. Mood (ANC)

- 1 = declarative
- 2 = interrogative
- 3 = imperative
- 4 = exclamative

7. Truncated ANC

- 1 = no
- 2 = yes

8. Clause type (ANC)

- 1 = simple
- 2 = complex-coord
- 3 = complex-sub
- 4 = not applicable

9. Type of Subject (ANC)

- 1 = noun
- 2 = pronoun
- 3 = null subject
- 4 = quantifier
- 5 = clause
- 6 = not applicable

10. Type of subject (IQT)

- 1 = null subject
- 2 = pronoun
- 3 = not applicable

11. IQT tense

- 1 = present
- 2 = continuous
- 3 = not applicable

12. ANC tense

- 1 = present
- 2 = past
- 3 = future
- 4 = perfect
- 5 = continuous
- 6 = gerund
- 7 = conditional
- 8 = imperative
- 9 = infinitive
- 10 = subjunctive
- 11 = not applicable

Bii. Dialogic

13. Response

1 = no

2 = other reply

3 = self reply

Biii. Generic

14. Genre

1 = Private dialogic discourse

2 = Public dialogic discourse

3 = Monologic discourse

Biv. Sociolinguistic

15. Sex

1 = male

2 = female

16. Age

European Portuguese English

1 = 18–25 18–25

2 = 26–40 26–45

3 = 41–60 46–65

4 = over 60 over 66

5 = unknown unknown

APPENDIX B. STATISTICAL ANALYSES

Table 9

Grammatical features of functional types of BEITQs.

	attitudinal (N = 45)	focusing (N = 12)	hortatory (N = 5)	informational (N = 75)	phatic (N = 4)	regulatory (N = 1)
Grammatical features						
IQT position						
medial	(–)	4 (33.3%)	(–)	(–)	(–)	(–)
final	45 (100.0%)	8 (66.7%)	5 (100.0%)	75 (100.0%)	4 (100.0%)	1 (100.0%)
Meta-polarity						
reversed	3 (6.7%)	(–)	(–)	16 (21.3%)	(–)	(–)
constant	40 (88.9%)	12 (100.0%)	4 (80.0%)	48 (64.0%)	4 (100.0%)	1 (100.0%)
neutral	2 (4.4%)	(–)	1 (20.0%)	11 (14.7%)	(–)	(–)
Polarity						
positive–negative	2 (4.4%)	(–)	(–)	12 (16.0%)	(–)	(–)
negative–positive	1 (2.2%)	(–)	(–)	4 (5.3%)	(–)	(–)
positive–positive	40 (88.9%)	12 (100.0%)	4 (80.0%)	48 (64.0%)	4 (100.0%)	1 (100.0%)
negative–negative	(–)	(–)	(–)	(–)	(–)	(–)
positive–neutral	1 (2.2%)	(–)	1 (20.0%)	9 (12.0%)	(–)	(–)
negative–neutral	1 (2.2%)	(–)	(–)	2 (2.7%)	(–)	(–)
Mood (ANC)						
declarative	43 (95.6%)	12 (100.0%)	(–)	64 (85.3%)	3 (75.0%)	1 (100.0%)
interrogative	1 (2.2%)	(–)	(–)	10 (13.3%)	(–)	(–)
imperative	1 (2.2%)	(–)	5 (100.0%)	(–)	1 (25.0%)	(–)
exclamative	(–)	(–)	(–)	1 (1.3%)	(–)	(–)

Truncated ANC						
no	45 (100.0%)	12 (100.0%)	5 (100.0%)	61 (81.3%)	4 (100.0%)	1 (100.0%)
yes	(–)	(–)	(–)	14 (18.7%)	(–)	(–)
Clause type (ANC)						
simple	18 (40.0%)	1 (8.3%)	4 (80.0%)	47 (62.7%)	1 (25.0%)	1 (100.0%)
complex-coord	4 (8.9%)	2 (16.7%)	(–)	4 (5.3%)	(–)	(–)
complex-sub	23 (51.1%)	9 (75.0%)	1 (20.0%)	24 (32.0%)	3 (75.0%)	(–)
not applicable	(–)	(–)	(–)	(–)	(–)	(–)
Type of Subject (ANC)						
noun	10 (22.2%)	4 (33.3%)	1 (20.0%)	7 (9.3%)	1 (25.0%)	(–)
pronoun	35 (77.8%)	8 (66.7%)	4 (80.0%)	53 (70.7%)	2 (50.0%)	1 (100.0%)
null subject	(–)	(–)	(–)	3 (4.0%)	(–)	(–)
quantifier	(–)	(–)	(–)	(–)	(–)	(–)
clause	(–)	(–)	(–)	2 (2.7%)	(–)	(–)
not applicable	(–)	(–)	(–)	10 (13.3%)	1 (25.0%)	(–)
Type of Subject (IQT)						
null subject	2 (4.4%)	(–)	(–)	(–)	(–)	(–)
pronoun	7 (15.6%)	5 (41.7%)	(–)	16 (21.3%)	(–)	1 (100.0%)
not applicable	36 (80.0%)	7 (58.3%)	5 (100.0%)	59 (78.7%)	4 (100.0%)	(–)
IQT Tense						
present	9 (20.0%)	5 (41.7%)	(–)	16 (21.3%)	(–)	1 (100.0%)
continuous	(–)	(–)	(–)	(–)	(–)	(–)
not applicable	36 (80.0%)	7 (58.3%)	5 (100.0%)	59 (78.7%)	4 (100.0%)	(–)
ANC Tense						
present	33 (73.3%)	8 (66.7%)	(–)	35 (46.7%)	2 (50.0%)	1 (100.0%)
past	6 (13.3%)	3 (25.0%)	(–)	11 (14.7%)	(–)	(–)
future	(–)	(–)	(–)	5 (6.7%)	(–)	(–)
perfect	3 (6.7%)	(–)	(–)	6 (8.0%)	(–)	(–)
continuous	2 (4.4%)	1 (8.3%)	(–)	5 (6.7%)	(–)	(–)
gerund	(–)	(–)	(–)	(–)	1 (25.0%)	(–)
conditional	(–)	(–)	(–)	2 (2.7%)	(–)	(–)
imperative	1 (2.2%)	(–)	5 (100.0%)	(–)	1 (25.0%)	(–)
infinitive	(–)	(–)	(–)	(–)	(–)	(–)
subjunctive	(–)	(–)	(–)	(–)	(–)	(–)
not applicable	(–)	(–)	(–)	(–)	(–)	(–)

Table 10

Grammatical features of functional types of EPITQs.

	attitudinal (N = 67)	challeng (N = 1)	facilit (N = 3)	focusing (N = 11)	hortat (N = 11)	informat (N = 91)	phatic (N = 51)	regulat (N = 162)n
Grammatical features								
IQT position								
medial	(–)	(–)	(–)	10 (90.9%)	(–)	3 (3.3%)	(–)	2 (1.2%)
final	67 (100.0%)	1 (100.0%)	3 (100.0%)	1 (9.1%)	11 (100.0%)	88 (96.7%)	51 (100.0%)	160 (98.8%)
Meta-polarity								
reversed	51 (76.1%)	(–)	3 (100.0%)	7 (63.6%)	4 (36.4%)	72 (79.1%)	7 (13.7%)	142 (87.7%)
constant	13 (19.4%)	1 (100.0%)	(–)	2 (18.2%)	7 (63.6%)	17 (18.6%)	43 (84.3%)	20 (12.3%)
neutral	3 (4.5%)	(–)	(–)	2 (18.2%)	(–)	2 (2.2%)	1 (2.0%)	(–)
Polarity								
positive-negative	49 (73.1%)	(–)	3 (100.0%)	7 (63.6%)	3 (27.3%)	72 (79.1%)	5 (9.8%)	142 (87.7%)
negative-positive	2 (3.0%)	(–)	(–)	(–)	1 (9.1%)	(–)	2 (3.9%)	(–)
positive-positive	(–)	(–)	(–)	1 (9.1%)	7 (63.6%)	4 (4.3%)	42 (82.4%)	(–)

negative-negative	13 (19.4%)	1 (100.0%)	(-)	1 (9.1%)	(-)	13 (14.3%)	1 (2.0%)	20 (12.3%)
positive-neutral	3 (4.5%)	(-)	(-)	2 (18.2%)	(-)	2 (2.2%)	(-)	(-)
negative-neutral	(-)	(-)	(-)	(-)	(-)	(-)	1 (2.0%)	(-)
Mood (ANC)								
declarative	64 (95.5%)	1 (100.0%)	3 (100.0%)	11 (100.0%)	4 (36.4%)	85 (93.4%)	51 (100.0%)	162 (100.0%)
interrogative	3 (4.5%)	(-)	(-)	(-)	(-)	6 (6.6%)	(-)	(-)
imperative	(-)	(-)	(-)	(-)	7 (63.6%)	(-)	(-)	(-)
exclamative	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)
Truncated ANC								
no	67 (100.0%)	1 (100.0%)	2 (66.7%)	10 (90.9%)	10 (90.9%)	87 (95.6%)	44 (86.3%)	161 (99.4%)
yes	(-)	(-)	1 (33.3%)	1 (9.1%)	1 (9.1%)	4 (4.4%)	7 (13.7%)	1 (0.6%)
Clause type								
simple	40 (59.7%)	1 (100.0%)	1 (33.3%)	3 (27.3%)	2 (18.2%)	56 (61.5%)	17 (33.3%)	96 (59.3%)
complex-coord	(-)	(-)	(-)	(-)	2 (18.2%)	2 (2.2%)	6 (11.8%)	6 (3.7%)
complex-sub	27 (40.3%)	(-)	1 (33.3%)	6 (54.5%)	6 (54.5%)	29 (31.9%)	18 (35.3%)	58 (35.8%)
not applicable	(0.0%)	(-)	1 (33.3%)	2 (18.2%)	1 (9.1%)	4 (4.4%)	10 (19.6%)	2 (1.2%)
Type of Subject (ANC)								
noun	15 (22.4%)	1 (100.0%)	(-)	4 (36.4%)	3 (27.3%)	19 (20.9%)	8 (15.7%)	38 (23.5%)
pronoun	10 (14.9%)	(-)	(-)	(-)	(-)	18 (19.8%)	5 (9.8%)	35 (21.6%)
null subject	33 (49.3%)	(-)	1 (33.3%)	5 (45.5%)	7 (63.6%)	47 (51.6%)	27 (52.9%)	79 (48.8%)
quantifier	(-)	(-)	(-)	(-)	(-)	1 (1.1%)	1 (2.0%)	(-)
clause	7 (10.4%)	(-)	(-)	(-)	(-)	(-)	(-)	4 (2.5%)
not applicable	2 (3.0%)	(-)	2 (66.7%)	2 (18.2%)	1 (9.1%)	6 (6.6%)	10 (19.6%)	6 (3.7%)
Type of Subject (IQT)								
null subject	2 (3.0%)	(-)	2 (66.7%)	1 (9.1%)	(-)	(-)	16 (31.4%)	(-)
pronoun	(-)	(-)	(-)	(-)	(-)	(-)	(-)	(-)
not applicable	65 (97.0%)	1 (100.0%)	1 (33.3%)	10 (90.9%)	11 (100.0%)	91 (100.0%)	35 (68.6%)	162 (100.0%)
IQT tense								
Present	61 (91.0%)	1 (100.0%)	3 (100.0%)	8 (72.7%)	11 (100.0%)	36 (39.6%)	6 (11.8%)	162 (100.0%)
continuous	(-)	(-)	(-)	1 (9.1%)	(-)	(-)	16 (31.4%)	(-)
not applicable	6 (9.0%)	(-)	(-)	2 (18.2%)	(-)	55 (60.4%)	29 (56.9%)	(-)
ANC tense								
present	8 (11.9%)	(-)	1 (33.3%)	(-)	3 (27.3%)	32 (35.2%)	13 (25.5%)	1 (0.6%)
past	44 (65.7%)	1 (100.0%)	(-)	8 (72.7%)	(-)	43 (47.3%)	21 (41.2%)	134 (82.7%)
future	5 (7.5%)	(-)	(-)	(-)	(-)	5 (5.5%)	4 (7.8%)	6 (3.7%)
perfect	(-)	(-)	(-)	(-)	(-)	1 (1.1%)	(-)	(-)
continuous	1 (1.5%)	(-)	(-)	(-)	(-)	(-)	1 (2.0%)	(-)
gerund	(-)	(-)	1 (33.3%)	(-)	(-)	(-)	(-)	3 (1.9%)
conditional	2 (3.0%)	(-)	(-)	(-)	(-)	2 (2.2%)	1 (2.0%)	1 (0.6%)
imperative	(-)	(-)	(-)	(-)	4 (36.4%)	(-)	(-)	(-)
infinitive	3 (4.5%)	(-)	(-)	(-)	(-)	3 (3.3%)	1 (2.0%)	9 (5.6%)
subjunctive	4 (6.0%)	(-)	(-)	1 (9.1%)	3 (27.3%)	1 (1.1%)	(-)	6 (3.7%)
not applicable	(-)	(-)	1 (33.3%)	2 (18.2%)	1 (9.1%)	4 (4.4%)	10 (19.6%)	2 (1.2%)

References

- Achiri-Taboh, B., 2015. A generalized question tag in English. *English Today* 31 (1), 48–54.
- Aijmer, K., 1979. The function of tag questions in English. In: Pettersson, T. (Ed.), *Papers from the Fifth Scandinavian Conference of Linguistics*. Acta Universitatis Lundensis and Almqvist and Wiksell, Lund and Stockholm, pp. 9–17.
- Aksu-Koç, A., Slobin, D.I., 1986. A psychological account of the development and use of evidentials in Turkish. In: Chafe, W., Nichols, J. (Eds.), *Evidentiality*. Ablex, Norwood, pp. 159–167.
- Algeo, J., 1990. It's a myth, *innit*? Politeness and the English tag questions. In: Ricks, C.h., Michaels, L. (Eds.), *The State of the Language*. Centennial Books. University of California Press, Berkeley, pp. 443–450.
- Algeo, J., 2006. *British or American English?*. Cambridge University Press Cambridge, UK.
- Allerton, D., 2009. Tag Questions. In: Rohdenburg, G., Schlüter, J. (Eds.), *One Language, Two Grammars: Differences between British and American English*. Cambridge University Press, pp. 306–323.
- Andersen, G., 1997. I goes you hang it up in your shower, innit? He goes yeah. The use and development of invariant tags and follow-ups in London teenage speech. Paper Given at First Language Variation Workshop, Reading, April 1997.

- Andersen, G., 2001. *Pragmatic Markers and Sociolinguistic Variation: A Relevance-Theoretic Approach to the Language of Adolescents. Pragmatics and Beyond*, vol. 84. John Benjamins, Amsterdam.
- Andersen, G., 1998. Are tag questions questions? Evidence from spoken data. In: Paper presented at the 19th ICAME Conference. Belfast, Ma, 1998. <http://gandalf.aksis.uib.no/~gisle/pdf/BELFAST.pdf> (accessed, 30 Dec 2020).
- Agresti, A., 2012. *Categorical Data Analysis*. Wiley, London.
- Gómez González, M.Á., Mackenzie, J.L., 2012. The question of tag questions in English and Spanish. In: Moskowich, I., Crespo, B. (Eds.), *Encoding the Past, Decoding the Future: Corpora in the 21st Century*. Cambridge Scholarly Press, Newcastle upon Tyne, pp. 59–97.
- Gómez González, M.Á., Dehé, N., 2014. Canonical tag questions in English, Spanish and Portuguese. A discourse-functional study. *Languages in Contrast* 14 (1), 93–126.
- Gómez González, M.Á., 2018. God that came out quick didn't it eh?: Variable and invariable tag questions in British English. In: Gómez González, M.Á., Mackenzie, J.L. (Eds.), *The Construction of Discourse as Verbal Interaction . Pragmatics & Beyond New Series*, vol. 296. Benjamins, Amsterdam, pp. 109–144.
- Gómez González, M.Á., Dehé, N., 2020. The pragmatics and prosody of variable tag questions in English: Uncovering function-to-form correlations. *Journal of Pragmatics* 158, 33–52.
- Axelsson, K., 2011a. A cross-linguistic study of grammatically dependent question tags: data and theoretical implications. *Studies in Language* 35 (4), 793–851. <https://doi.org/10.1075/sl.35.4.02axe>.
- Axelsson, K., 2011b. *Tag Questions in Fiction Dialogues* [Doctoral dissertation]. University of Gothenburg Press.
- Axelsson, K., 2014. Confirmation-demanding tag questions in fiction dialogue. In: Glynn, D., Sjölin, M. (Eds.), *Subjectivity and Epistemicity: Corpus, Discourse, and Literary Approaches to Stance*. Lund Studies in English, vol. 117, pp. 189–209.
- Barbosa, P., Santos, P., Veloso, R., 2020. Tipos de frase e força ilocutória. In: Raposo, E.P., Bacelar, M.F., Mota, M.A.C., Segura, L., Mendes, A., Andrade, A. (Eds.), *Gramática do Português*, vol. III. Fundação Calouste Gulbenkian, pp. 2535–2563.
- Barron, A., Pandavara, I., Muderacj, K., 2015. Tag questions across Irish English and British English: A corpus analysis of form and function. *Multilingua* 34, 495–525.
- Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc.: Ser. B (Methodol.)* 57 (1), 289–300.
- Berland, U., 1997. *Invariant Tags: Pragmatic Functions of innit, okay, right and yeah in London Teenage Conversation*. University of Bergen, Unpublished MA thesis.
- Berry, M., 1981a. Systemic linguistics and discourse analysis: a multi-layered approach to exchange structure. In: Coulthard, M., Montgomery, M. (Eds.), *Studies in Discourse Analysis*. Routledge & Kegan Paul, London, pp. 120–145.
- Berry, M., 1981b. Towards layers of exchange structure for directive exchanges. *Network* 2, 23–32.
- Berry, M., 2016. Dynamism in exchange structure. *Engl. Text Constr.* 9 (1), 33–55. <https://doi.org/10.1075/etc.9.1.03ber>.
- Biber, D., Egbert, J., 2016. Register Variation on the Searchable Web: A Multi-Dimensional Analysis. *Journal of English Linguistics* 44 (2), 95–137.
- Biber, D., Egbert, J., Zhang, M., 2018. Lexis and grammar as complementary discourse systems for expressing stance and evaluation. In: Mackenzie, J.L., Author (Eds.), *The Construction of Discourse as Verbal Interaction*. John Benjamins, Amsterdam, pp. 197–222.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., 1999. *Longman Grammar of Spoken and Written English*. Longman, London.
- Biscetti, S., 2006. Tag questions in courtroom discourse. In: Flowerdew, J., Gotti, M. (Eds.), *Studies in Specialized Discourse*. Peter Lang, Bern, pp. 209–238.
- Bonsignori, V., 2013. *English Tags: A Close-Up on Film Language, Dubbing and Conversation*. Cambridge Scholars Publishing, Newcastle upon Tyne.
- Boye, K., 2012. *Epistemic Meaning. A Cross-Linguistic and Functional-Cognitive Study*. De Gruyter Mouton, Berlin/Boston.
- Brazil, D., 2008. Tag questions. *Ilha do Desterro* 5.
- Bublitz, W., 1978. *Ausdrucksweisen der Sprechereinstellung im Deutschen und Englischen*. Untersuchungen zur Syntax, Semantik und Pragmatik der deutschen Modalpartikeln und Vergewisserungsfragen und ihrer englischen Entsprechungen, vol. 57. Walter de Gruyter, Tübingen.
- Calnan, A., Davidson, M., 1998. The impact of gender and its interaction with role and status on the use of tag questions in meetings. *Women Manage. Rev.* 13, 19–36.
- Cameron, D., McAlinden, F., O'Leary, K., 1988. Lakoff in context: The social and linguistic functions of tag questions. In: Coates, J., Cameron, D. (Eds.), *Women in Their Speech Communities: New Perspectives on Language and Sex*. Longman, London, pp. 74–93.
- Carvalho, A.M., Kern, J., 2019. The permeability of tag questions in a language contact situation. *Pragmatics* 29 (4), 463–492.
- Carter, R., McCarthy, M.J., 2006. *Cambridge Grammar of English: A Comprehensive Guide*. Spoken and Written English Grammar and Usage. Cambridge University Press, Cambridge.
- Cheng, W., Warren, M., 2001. 'She knows more about Hong Kong than you do isn't it': Tags in Hong Kong conversational English. *J. Pragm.* 33 (9), 1419–1439.
- Columbus, G., 2010. A comparative analysis of invariant tags in three varieties of English. *English World-Wide* 31, 288–310.
- Cortés, L., Camacho, M., 2005. *Unidades de segmentación y marcadores del discurso*. Arco Libros, Madrid.

- Criado Peña, M., 2016. 'They are going tomorrow, isn't it?' On the use of tag questions in Indian English and Hong Kong English. In: Moreno Ortiz, A., Pérez-Hernández, C. (Eds.), *EPiC Series in Language and Linguistics*, vol. 1, pp. 71–78.
- Cruz-Ferreira, M., 1981. Tag questions in Portuguese: grammar and intonation. *Phonetica* 38, 341–352.
- Cuenca, M.-J., 1997. Form-use mappings in tag questions. In: Liebert, W.-A., Redeker, G., Waugh, L.R. (Eds.), *Discourse and Perspective in Cognitive Linguistics*. Benjamins, Amsterdam, pp. 3173–3237.
- DeLancey, S., 2001. The mirative and evidentiality. *Journal of Pragmatics* 33 (3), 369–382.
- Dehé, N., Braun, B., 2013. The prosody of question tags in English. *English Language and Linguistics* 17 (1), 129–156.
- Duarte, L.C., 1985. Tag Questions in English and Portuguese: A Contextual Analysis [Unpublished Masters' thesis]. University of Florianópolis.
- Dubois, B.L., Crouch, I., 1975. The question of tag questions in women's speech: They don't really use more of them, do they". *Language in Society* 4 (3), 289–294.
- Eckert, P., McConnell-Ginet, S., 2003. *Language and Gender*. Cambridge University Press, Cambridge.
- Fischer, K., 2006. Towards an understanding of the spectrum of approaches to discourse particles: introduction to the volume. In: Fischer, K. (Ed.), *Approaches to Discourse Particles*. Elsevier, Amsterdam, pp. 1–12.
- Ford, C.E., Fox, B., Thompson, S., 2003. Social Interaction and grammar. In: Tomasello, M. (Ed.), *The New Psychology of Language Cognitive and Functional Approaches to Language Structure*, vol. 2. Lawrence Erlbaum Associates, New Jersey, pp. 119–144.
- Fraser, B., 2013. Combinations of Contrastive Discourse Markers in English. *International Review of Pragmatics* 5, 318–340.
- Gonçalves, F., 1995. *Negação Frásica em Português: Caracterização com Referência ao Processo de Aquisição*. Dissertação de Mestrado. FLUL, Lisboa.
- Govindan, I., Pillai, S., 2009. English question forms used by young Malaysian Indians. *Engl. Teacher* 38, 74–94.
- Heritage, J., Stivers, T., Mondada, L., Steensig, J., 2011. Territories of knowledge, territories of experience: empathic moments in interaction. In: Stivers, T., Mondada, L., Steensig, J. (Eds.), *The Morality of Knowledge in Conversation*. Cambridge University Press, Cambridge, pp. 159–183. <https://doi.org/10.1017/CBO9780511921674.008>.
- Hoffmann, S., Blass, A.K., Mukherjee, J., 2014. Canonical tag questions in Asian Englishes: forms, functions and frequencies in Hongkong English, Indian English and Singapore English [Paper presentation]. In: ICAME 32. The 32nd Conference of the International Computer Archive of Modern and Medieval English), June 1–5, Oslo, Norway.
- Holmes, J., 1983. The Functions of Tag Questions. *English Language Research Journal* 3, 40–65.
- Holmes, J., 1984. Hedging your Bets and Sitting on the Fence: Some Evidence for Hedges as Support Structures. *Te Reo* 27, 47–62.
- Holmes, J., 1995. *Women, Men and Politeness*. Longman, London.
- Holmes, J., 2001. *An Introduction to Sociolinguistics*. Pearson Education Limited, Essex, London.
- Huddleston, R., Pullum, G.K. (Eds.), 2002. *The Cambridge Grammar of the English Language*. Cambridge University Press, Cambridge.
- Hudson, R., 1975. The Meaning of Questions. *Language* 51 (1), 1–31.
- Hyland, K., 2005a. Stance and engagement: a model of interaction in academic discourse. *Discour. Stud.* 7 (2), 173–192.
- Hyland, K., 2005b. *Metadiscourse*. Continuum, London.
- Jespersen, O., 1924. *The Philosophy of Grammar*. George Allen & Unwin Ltd, London.
- Kim, J.B., Ahn, Y.A., 2008. English tag questions: corpus findings and theoretical implications. *English Language and Linguistics* 25, 103–126.
- Kimps, D., 2007. Declarative constant polarity tag Questions: A data-driven analysis of their form, meaning and attitudinal uses". *Journal of Pragmatics* 39 (2), 270–291.
- Kimps, D., 2016. English variable tag questions: A typology of their interpersonal meanings. University of Leuven: Department of Linguistics. Unpublished doctoral dissertation.
- Kimps, D., 2018. Tag questions in conversation: A typology of their interactional and stance meanings. Benjamins, Amsterdam.
- Kimps, D., Davidse, K., Cornillie, B., 2014a. A speech function analysis of British English tag questions in spontaneous dialogue. *Journal of Pragmatics* 66, 64–85.
- Kimps, D., Davidse, K., Cornillie, B., 2014b. The speech functions of tag questions and their properties. A comparison of their distribution in COLT and LLC. In: Davidse, K., Gentens, C., Ghesquière, L., Vandelanotte, L. (Eds.), *Corpus Interrogation and Grammatical Patterns*. Benjamins, Amsterdam, pp. 321–350.
- Kursa, M.B., Rudnicki, W.R., 2010a. Feature Selection with the Boruta Package. *Journal of Statistical Software* 36, 1–13.
- Labov, W., 1972. The Study of Language in its Social Context. In: Giglioli, P. (Ed.), *Language and Social Context*. Penguin, Harmondsworth, pp. 283–307.
- Labov, W., Fanshel, D., 1977. *Therapeutic Discourse: Psychotherapy as Conversation*. Academic Press, New York.
- Lakoff, R., 1975. *Language and Woman's Place*. Harper Colophon Books, New York.
- Landis, R.J., Koch, G.G., 1977. The measurement of observer agreement for categorical data. *Biometrics* 33 (1), 159–174.
- Lapadat, J., Seesahai, M., 1977. Male versus female codes in informal contexts. *Sociolinguistics Newsletter* 8, 7–8.
- Lyons, J., 1977. *Semantics*. Cambridge University Press, Cambridge.
- Martin, J.R., 1992. *English Text: System and Structure*. John Benjamins, Amsterdam.

- Martínez Caro, E., 2020. Discourse functions of question tags: Exploring the right periphery in English. In: Achiri-Taboh, B. (Ed.), *Exoticism in English Tag Questions (Strengthening Arguments and Caressing the Social Wheel)*. Cambridge Scholars Publishing, Newcastle-upon-Tyne, pp. 215–235.
- Martins, A.M., 2006. Emphatic affirmation and polarity: Contrasting European Portuguese with Brazilian Portuguese, Spanish, Catalan and Galician. In: Doetjes, J., González, P. (Eds.), *Romance Languages and Linguistic Theory. Selected papers from Going Romance 2004*. John Benjamins, Amsterdam, pp. 197–223.
- Martins, A.M., 2013. Emphatic polarity in European Portuguese and beyond. *Lingua* 128, 95–123.
- McGregor, W., 1995. The English 'tag question': A new analysis, is(n't) it? In: Hasan, R., Fries, P.H. (Eds.), *On Subject and Theme. A Discourse Functional Perspective*. Current Issues in Linguistic Theory. Benjamins, Amsterdam, pp. 91–121.
- McGregor, W., 1997. *Semiotic Grammar*. Clarendon, Oxford.
- Meyer, D., Zeileis, A., Hornik, K., 2020. *vcd: Visualizing Categorical Data*. R package version 1.4-7.
- Kursa, M.B., Rudnicki, W.R., 2010b. Feature selection with the boruta package. *J. Stat. Softw.* 36 (11), 1–13.
- Mithun, M., 2012. Tags: cross-linguistic diversity and commonality. *Journal of Pragmatics* 44 (15), 2165–2182.
- Moniz, H., Batista, F., Trancoso, I., Mata, A.I., 2011. Analysis of interrogatives in different domains. In: Esposito, A., Esposito, A.M., Martone, R., Müller, V.C., Scarpetta, G. (Eds.), *Toward autonomous, adaptive, and context-aware multimodal interfaces. Theoretical and practical issues*. Springer, Berlin, pp. 134–146.
- Moniz, H., Batista, F., Mata, A.I., Trancoso, I., 2016. Towards automatic language processing and intonational labeling in European Portuguese. In: Armstrong, M.E., Henriksen, N., Vanrell, M.M. (Eds.), *Intonational Grammar in Ibero-Romance: Approaches across Linguistic Subfields*. John Benjamins, Amsterdam, pp. 227–250.
- Moore, E., Podesva, R., 2009. Style, indexicality, and the social meaning of tag questions. *Language in Society* 38 (4), 283–311.
- Nasslin, S., 1984. *The English Tag Question: A Study of Sentences Containing Tags of the Type Isn't it?, Is it? Almquist, and Wiksell International, Nässlin*.
- Nelson, G., Wallis, S., Aarts, B., 2002. *Exploring Natural Language: The British Component of the International Corpus of English*. John Benjamins, Amsterdam.
- Norrick, N., 1995. Hunh-tags and evidentiality in conversation. *Journal of Pragmatics* 23, 687–692.
- Pearson, R. (2020). GoodmanKruskal: Association Analysis for Categorical Variables. R package version 0.0.3. <https://CRAN.R-project.org/package=GoodmanKruskal>.
- Pichler, H., 2013. *The Structure of Discourse-Pragmatic Variation*. Studies in Language Variation 13. John Benjamins, Amsterdam.
- Pons Bordería, S., 1998. Conexión y conectores. Estudio de su relación en el registro informal de la lengua. Anejo XXVII de Cuadernos de Filología. Universitat de Valencia, Valencia.
- Portolés, J., 1998. *Marcadores del discurso*. Ariel, Barcelona.
- Quirk, R., Greenbaum, S., Leech, G., Svartvik, J., 1985. *A Comprehensive Grammar of the English Language*. Longman, London.
- R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria <https://www.R-project.org/>.
- Recsky, L., 2006. Epistemic modality and spoken discourse: an English-Portuguese cross-linguistic investigation. *Ling. Ensino* 9 (1), 159–185.
- Rodrigues, I., 1999. In: *Turn-maintaining signals in spoken Portuguese*. University of Porto Press, pp. 723–731.
- Rodríguez Muñoz, F.J., 2009. Estudio sobre las funciones pragmadiscursivas de ¿eh? ¿no? en el español hablado. *Revista de lingüística teórica y aplicada* 47 (1), 83–101.
- Santos, A.L., 2003. The acquisition of answers to yes/no questions in Portuguese: syntactic, discourse and pragmatic factors. *Journal of Portuguese Linguistics* 2 (1), 61–91.
- Siegel, S., Castellan, J.N., 1988. *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, New York.
- Simon-Vandenberghe, A.-M., Aijmer, K., 2007. *The Semantic Field of Modal Certainty. A Corpus-Based Study of English Adverbs*. Mouton de Gruyter, Berlin.
- Stenström, A.-B., 1997. Tags in teenage talk". In: Fries, U., Muller, V., Schneider, P. (Eds.), *Studies in English corpus linguistics*. Rodopi, Amsterdam, pp. 139–147.
- Stenström, A.-B., 2005. Teenagers' tags in London and Madrid. In: McCafferty, K., Bull, T., Killie, K. (Eds.), *Contexts – Historical, Social, Linguistic: Studies in Celebration of Toril Swan*. Peter Lang, Bern, pp. 279–291.
- Stenström, A.-B., Andersen, G., Hasund, I.K., 2002. *Trends in Teenage Talk: Corpus Compilation, Analysis and Findings*. John Benjamins, Amsterdam.
- Tagliamonte, S.A., 2016. *Teen Talk: The Language of Adolescents*. Cambridge University Press, Cambridge.
- Takahashi, M., 2014. A comparative study of tag questions in four Asian Englishes from a corpus-based approach. *Asian Englishes* 16 (2), 101–124.
- Thompson, G., 2001. Interaction in academic writing: Learning to argue with the reader. *Applied Linguistics* 22 (1), 58–78.
- Tomaselli, M.V., Gatt, A., 2015. Italian Tag Questions and their Conversational Functions. *Journal of Pragmatics* 84, 54–82.
- Torgersen, E., Gabrielatos, C., Hoffmann, S., Fox, S., 2011. A Corpus-based study of pragmatic markers in London English. *Corpus Linguistics and Linguistic Theory* 7 (1), 93–118.
- Tottie, G., Hoffmann, S., 2006. Tag questions in British and American English. *Journal of English Linguistics* 34, 283–311.
- Tottie, G., Hoffmann, S., 2009. Tag questions in English: The first century. *Journal of English Linguistics* 37, 130–161.