A constructionalist account of *why*-fragments and Mad Magazine sentences: the 'Sceptical Small' construction[1]

YOLANDA FERNÁNDEZ-PENA ⬡ and
JAVIER PÉREZ-GUERRA ⬡
Universidade de Vigo
(Received 4 October 2023; revised 17 May 2024)

This article focuses on two fragmentary constructions in English: *why*-fragments (WFs), such as *Why (deal with) why-fragments?*, and Mad Magazine sentences (MMs), such as *(Me) paint the house purple?* While both types can be equivalent in meaning to their corresponding fully fledged interrogative sentences, they can also be used to convey a specific nuance of scepticism regarding a particular proposition. To explore the specific nuance enriching the canonical interpretation (i.e. equivalent to that of the corresponding complete questions) of WFs and MMs, and their potential constructionalisation in contemporary English, two corpus-based studies were conducted using data from the BNC1994 DS, Spoken BNC2014 and COCA. The results show that MMs seem to be fully constructionalised, while the significant trends attested for WFs indicate an ongoing process of constructionalisation, at least in contemporary British English. The evidence also shows that both may be classed as examples of an umbrella 'Sceptical Small' construction.

Keywords: fragment, *why*-fragment, Mad Magazine, construction, Construction Grammar

1   Why investigate *why*-fragments and Mad Magazine sentences?

This study focuses on two fragmentary patterns: *why*-fragments (WFs) and Mad Magazine constructions (MMs, after Akmajian 1984), illustrated in (1) and (2), respectively.

(1) [– I was nasty when I replied to the editor.]
       – **Why nasty?**
(2) [– You were nasty when you replied to the editor.]
       – **You nasty?**

'Fragments' are defined as functionally stand-alone, syntactically and prosodically independent expressions that are formally reduced but nonetheless semantically,

discursively and pragmatically equivalent to the corresponding complete sentences in propositional meaning, force and communicative function (Fernández-Pena 2021). WFs and MMs are classed as fragments because they lack a fully fledged clausal design, are recognised as independent syntactic units, have an independent intonational realisation and, in terms of meaning, force and pragmatic potential, behave in the same way as orthodox, canonical complete sentences in discourse.

WFs are introduced by *why*, and followed by so-called 'remnants', which may be nominal (3), adjectival (4), adverbial (5), prepositional (6) or bare infinitival predicates (7):

(3) Why ***why*-fragments**?
(4) Why **so specific**?
(5) Why **now**?
(6) Why **in this paper**?
(7) Why **deal with *why*-fragments**?

Like other fragments, WFs can convey the 'orthodox' or 'canonical' interrogative interpretation of a complete *why*-question, as in (8a), and/or a specific 'unique' or 'modal' interpretation (Gordon & Lakoff 1971: 481; Johnson 1975: 487; Huddleston & Pullum *et al.* 2002: 835, 874; Weir 2017: 406; Zaitsu 2018, 2020) involving subjective scepticism directed at a previous proposition, as illustrated in (8b):

(8) [– Laura deals with *why*-fragments in her dissertation.]
      – Why deal with *why*-fragments in a dissertation?

  (a) Why did Laura deal with *why*-fragments in her dissertation?
  (b) Why should Laura deal with *why*-fragments in her dissertation? What's the reason for it? Why not in her opening lecture?

MMs consist of two compulsory components: 'MM1', which resembles a subject of the construction; and 'MM2', a predicate-like member which may mirror a predicative complement (as in (2) above), an object (9), a prepositional complement (10) or an adjunct (11).

(9) [– I also discussed MMs in my presentation.]
      – You, **MMs**?
(10) [– I also dealt with MMs in my presentation.]
       – You, **with MMs**?
(11) [– They eventually published my paper in *ELL*.]
       – Them **in *ELL***?

Like WFs, MMs can convey the interpretation corresponding to a canonical complete question ('Did you discuss MMs in your presentation', in (9)) and/or have a construction-specific nuance ('Did ʏᴏᴜ discuss MMs in your presentation? I am very sceptical about that…'), according to which the possibility of the action, state or event in general is placed in question (Donaldson 2013: 13).

This study reports on an analysis of both construction types in contemporary English. Section 2 justifies the use of a Construction Grammar approach, and the analysis of WFs and MMs as examples of an umbrella Sceptical Small construction in which form and meaning are paired in a specific, non-compositional way (Goldberg 2019: 73). The same section contains a review of the relevant literature, and argues that WFs and MMs should not be classed as expressions resulting from the application of deletion or ellipsis operations to complete questions. Section 3 analyses WFs and MMs in corpora of contemporary English in order to examine the recent constructionalisation processes in each. Section 3.1 focuses on WFs in representative samples retrieved from the spoken components of the two editions of the *British National Corpus*. The aim of this analysis is twofold: firstly, to determine whether the specific interpretation of WFs can be motivated by linguistic factors, such as the category of the remnant and the existence of categorial mismatches between the WF and the previous linguistic context; and, secondly, to explore the evolution of that conventionalised meaning in contemporary English. The null hypothesis is the lack of connection between the specific non-compositional interpretation of WFs and their linguistic characterisation. Section 3.2 focuses on MMs in contemporary American English, using data retrieved from the *Corpus of Contemporary American English*. The analysis explores formal factors, such as the category and syntactic function of the second member of the MM and the distribution of valid examples across text types and registers, and potential meaning-related determinants, such as polarity, the referential status of MM2 and the subjective reference of the MM's pronominal first member, in order to determine whether the MM pattern has been subjected to constructional processes in recent times. Section 4 reports on the main findings and proposes a characterisation of WFs and MMs from a (Cognitive) Construction Grammar perspective. Finally, section 5 contains a summary, concluding remarks and avenues for further research.

2 *Why*-fragments and Mad Magazine sentences as constructions

From a Construction Grammar perspective, a construction is defined as a 'syntactic pattern dedicated to semantic and pragmatic purposes not knowable from its form alone' (Fillmore *et al.* 1988: 505). In other words,

> whenever it is not possible to predict all of the facts about the use, internal composition, combinatory potential, or meaning of the pattern under study to some independently motivated principles or already known construction(s), it may become necessary to propose a separate construction. (Boas 2013: 235)

Figure 1, from Croft (2001: 18), illustrates the direct link between form and meaning in a construction.

The present study is based on the categorisation of patterns such as (1) and (2) above as constructions, i.e. as 'a formal type in its own right … a FORMAL IDIOM' (Lambrecht 1990: 218). The definition of WFs and MMs as constructions may be justified, firstly, by the fact that both have specific formal properties:
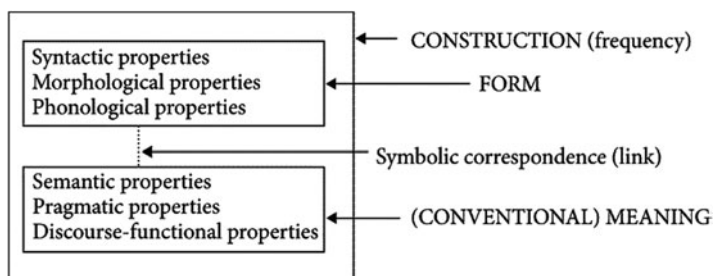
Figure 1. Defining properties of a construction

- WFs and MMs necessarily have an interrogative (or exclamatory, in non-prototypical instances) (non-assertive) illocutionary force, which is manifested through non-declarative intonation – Akmajian (1984: 4): 'MMs can be used to exclaim, but not to inform (i.e. assert)'; Progovac (2006: 33): MMs 'can be interpreted as anything but an assertion, ranging over exclamative, imperative, and other possibilities'.
- When compared to prototypical full sentences, WFs and MMs are reduced or 'small' constructions that lack tensed verbal operators (Tense, Aux, Modal) – see Pskit (2017).
- WFs and MMs have a comparable structural design, as shown in (12) and (13), respectively:

(12) WFs: *why* (NP) {V, N, A (, P, Adv)}
(13) MMs: (NP) {V, N, A (, P, Adv?)}[2] (adapted from Akmajian 1984)

- Subjects and nominals are optional: rare in WFs (see (14)–(16)) and pervasive in MMs (see (17)):

(14) why **you** bored? (BNC 1994 DS: KBE 3475)
(15) Why **you** a bit wet? (BNC1994 DS: KD1 3255)
(16) Why **you** in there? (BNC1994 DS: KDE 2341)
(17) (**Me**) get a respectable job? (Lambrecht 1990)

If explicit (and pronominal), possibly due to the fact that the construction is by definition nonfinite (Akmajian 1984; Lambrecht 1990), subjects are accusative (*Why me*$_{ACC}$*?*, *Me*$_{ACC}$ *worry?*) and necessarily referential (*Why there$_{dummy}$*?*, **It rain again?*).

Secondly, WFs and MMs can convey a canonical interrogative interpretation but also a specific (or 'inferred'; De Vaere *et al.* 2020) meaning that is not precisely identical to the interpretation encoded by the corresponding complete-clause/sentence constructions. Both patterns may thus be classed as unique pairings of form and meaning in their own right: in other words, as constructions. The corpus-based studies on WFs and MMs in section 3 reveal that some patterns have only a canonical question meaning, others are often interpreted with a 'specific' (enriched) nuance, and others have both meanings, in

---

[2] 'V', 'N', 'A', 'P' and 'Adv' represent lexical and phrasal projections.

keeping with (Cognitive) Construction Grammar premises. One interpretation may be privileged over the other in pairs of related or competing fragmentary and complete expressions, as observed by Bauer & Hoffman (2020: 245) when they hypothesise that

> there will be contexts in which one construction strongly preempts the other alternative, which in effect also minimises constructional synonymy. This could, e.g., lead to a situation where 'full forms' and 'reduced fragments' … acquire different semantic restrictions.

As regards synonymy in constructions, Goldberg (2013: 19) contends that '[i]t is possible for a construction to be ambiguous: the same form being paired with unrelated functions, just like ambiguous words'. To illustrate the plausibility of specific interpretations which are not necessarily identical to the orthodox meaning conveyed by the corresponding complete questions, let us consider examples (1) and (2) above, repeated here as (18) and (19) for convenience:

(18) Why nasty? (WF)
    'Why were you nasty?' (= complete question)
    'Why did you have to / should you be nasty? Were there reasons for your behaviour?' (specific interpretation)
(19) You nasty? (MM)
    'Were you nasty?' (= complete question)
    'You were nasty? I can't believe it!' (specific interpretation)

Two considerations are in order here as regards the specific interpretation of WFs and MMs not directly derivable from the meanings conveyed by their constituents. Firstly, since WF and MM expressions are, by definition, nonfinite or untensed, the specific interpretation is never directed at a situation, event, action or fact but, as Lambrecht (1990: 220; our italics) explains, at 'a *proposition* which was expressed (or contextually implied) in the immediately preceding discourse'. To quote Donaldson (2013: 13), both constructions, therefore, 'serve to isolate the action from a specific time and question the possibility of such an action in general'. The examples in (20) and (21) reflect the bare propositions implied by the specific interpretation of WF and MM constructions:

(20) WFs: Why give me a rise in two weeks' time?:
    [$_{proposition}$ X give me a rise in two weeks' time]
    Why a rise?:
    [$_{proposition}$ A rise be/exist]
    Why in two weeks' time?:
    [$_{proposition}$ X be/exist in two weeks' time]
(21) MMs: My boss give me a rise?:
    [$_{proposition}$ My boss give me a rise]
    My boss a generous man?:

[$_{proposition}$ My boss be a generous man]
My boss generous?:
[$_{proposition}$ My boss be generous]

Secondly, according to Donaldson's (2013: 6) analysis of elliptical *why*-questions, both constructions encourage 'the introduction of skepticism' towards the proposition implied by the WF/MM. From a different perspective, Akmajian (1984: 2; our italics) claims in his study on MMs that the constructions 'are used by speakers to express surprise, disbelief, *skepticism*, scorn, and so on, at some situation or event'. Taking this a step further, the specific effects of modality, surprise, disbelief and/or scorn conveyed by WF/MM constructions may be interpreted as (pragmatic, communicative) by-products of propositional scepticism. By stripping a tensed question and keeping only its non-verbal components (sometimes even removing the predicate's subject/agent), WFs and MMs focus on and challenge the question's non-deictic propositional skeleton, thus paving the way for scepticism.[3]

These two types of what we will call the umbrella 'Sceptical Small' schema are sometimes analysed as reduced expressions resulting from the application of deletion or ellipsis to full questions (for deletion approaches, see Johnson 1975; Nakao *et al.* 2012; Weir 2014; Yoshida *et al.* 2015 and Alves-Castro 2020; for accounts not based on reconstruction, see Kim 2017; Kim & Abeillé 2019; for base-generation proposals, see Progovac 2006). The dismissal of a deletion/ellipsis interpretation of WFs and MMs is amply supported in the literature, as the following short summary of the arguments shows:

- WFs and MMs may be exophoric or discourse-new and even discourse-initial (see Donaldson 2013: 29). In other words, they may lack linguistic correlates, as in (22) and (23). This circumstance clearly invalidates an ellipsis-based account (except in infrequent syntactic operations, such as deletion via Merchant's 2004 'innocent' reconstruction, for example).

(22) [– But where do you find food and stuff? Food? Food.]
       – **You hungry!** (COCA:2008:TV My Name Is Earl)
(23) [– Her pastor was leaving town.]
       – **Why the rush?** (Kim 2017: 741)

- Referring constituents in WFs and MMs may exhibit semantic and/or grammatical replacement, as in (24)–(26). In (24) *at that age* in the WF co-refers and replaces *when you were 9 years old*. In (25) the adjective *worried* in the MM is connected to the verb *worry*. Finally, accusative *him* resumes the nominative nominal *Saddam* in (26). Such replacement operations cannot be accounted for by deletion.

---

[3] Siemund (2018: §10.1.5) provides another example of a sceptical (or, in his words, 'incredulity response') construction, illustrated by *Do this thing – never!*, also headed by a nonfinite verb.

(24) [– That's what drew you into the movement when you were 9 years old.]
      – **Why at that age?** (Kim & Abeillé 2019: 380)

(25) [– I didn't want to worry you.]
      – **Me, worried?** (COCA:1997:TV Nash Bridges)

(26) They say that Saddam is a dictator, but **why him?** (Kim & Abeillé 2019: 978)

• WFs and MMs tolerate violations of connected prepositions, as in the WF example in (27). This linguistic reduction in the WF can hardly be accounted for by orthodox deletion.

(27) [– Robin's body had been found **at** the golf course just a half mile from her home under some leaves.]
      – Why **on** the golf course? (COCA:2005:SPOK NBC_Dateline)

• The so-called remnant in the WF (our data do not contain equivalent examples for MMs) may correlate with an item attested in a complex syntactic environment in the antecedent clause or 'island' (e.g. coordinating construction (28) or modifier/complement of a noun phrase (29)). This would be ungrammatical if the WF were an instance of ellipsis, since extraction from an elliptical context would lead to a movement violation.[4]

(28) [– Well, I think what's happening is <u>both risky and immoral</u>.]
      – **Why immoral?** (COCA:2009:SPOK PBS_NewsHour) (Coordinate Structure Constraint)

(29) [– It's got to be <u>someone with a with a caretaker</u> hasn't it be elect them, there isn't one is there?]
      – **Why a caretaker?** (BNC1994 DS: KBW 15592) (Complex NP Constraint)

Further evidence for the constructional characterisation of WFs and MMs in English, and their analysis as Sceptical Small constructions is provided by other interrogative[5] *wh*-fragmentary constructions which do not comply with the defining features of the Sceptical Small construction:

• 'reactive' *what*-X construction: As a special pairing of form and meaning, this construction, illustrated in [– *We can narrow down who takes those.*] – *What the tall ones* (Põldvere & Paradis 2019: 66), conveys meaning that is not compositional since it 'expresses a reaction and a request at the same time' (p. 65) and has a form which can be regarded as 'peculiar in that *what* is directly connected to a phrasal or clausal complement in a way that does not agree with traditional phrase and clause structure

---

[4] As pointed out by a reviewer, ellipsis strategies differ substantially in terms of island sensitivity. As reported in Yoshida *et al.* (2015: 342–3), while island violation is deemed unacceptable in stripping, both sluicing and *why*-stripping tolerate extractions from complex syntactic environments.

[5] On *what* and *how* exclamative fragments, see Heine *et al.* (2020) and Siemund (2015, 2017). On *what with* (declarative) absolute constructions (e.g. *What with him being a physician,* [*he's more pragmatic*]), see Kortmann (1991: 202) and Kim & Davies (2019).

rules' (p. 66) and 'functions as a parenthetical' (Põldvere & Paradis 2020: 318). Since *what* is a mandatory member in the reactive construction, the latter cannot be included within the category of WF.

- split interrogative: As claimed by Michaelis & Feng (2015: 15), expressions such as *What are you, a senior?* contain 'several idiosyncratic properties that suggest [they are] more than the sum of [their] parts', and should, therefore, be classed as constructions. Since the *wh*-clauses in split interrogatives are finite, these constructions deviate from the defining formal schema of WFs.

- WHX constructions, such as *what the hell/fuck are you talking about?*: The constructional analysis of these expressions is appropriate, according to Hugou (2017), because 'the meaning of the WHX construction is justified by the meaning of its parts, but is not computable from them' (p. 4). These constructions express 'intensity …, an unmarked state, a reference value… and a marked state' (p. 3). Since the *wh*-clause is finite, WHX constructions cannot be regarded as a subtype of the WF construction.

- Siemund's (2018: §10.1.1) minor clauses introduced by *wh*-forms, such as *how come, what if, how/what about* [*ing*|nominal]: These expressions do not fall within the scope of WFs because they are either finite or introduced by a *wh*-form other than *why*.

- parenthetical *what* in *You spend what 17,000 pounds on one of these* (Dehé & Kavalova 2006): Parenthetical *what* is excluded from the WF category because the construction is built exclusively upon *what*.

- sluice expressions such as *what?, who?* and Woods & Vicente's (2021) 'metacommunicative-*why* fragments' such as *why?*: Sluice expressions imply polar questioning, i.e. a canonical interrogative interpretation, in addition to a 'process of utterance coercion or accommodation …, triggered by the inability to ground the previous utterance' (Fernández *et al.* 2007: 404). The fixed atomic (one-word) design of sluice structures makes them uninteresting from a constructional variationist perspective and justifies their exclusion from the category of (prototypical) WFs.

- 'massive pied-piping' constructions such as *run where?* (in Weir 2017: 403): These structures are similar to regular sluicing expressions (e.g. *John ate something but I don't know what*) but lack additional clausal material (e.g. the subject in *run where?*) Even though they deserve constructional treatment on the grounds that they convey the interpretation of canonical questions as well as the specific nuance 'Where should I run / Where am I supposed to run' (Weir 2017: 406), massive pied-piping patterns are construed around *wh*-forms other than *why* and, in consequence, cannot be considered WF constructions.

- sluice stranding: This is defined by Culicover (2013: 109) as 'sluicing in which the *wh*-phrase precedes a preposition', as in *what about/with/for?* or *who with?* Since the sluicing stranding of *why* is not acceptable, these examples of sluice stranding cannot be classed as *why*-fragment constructions.

This section has set the stage for the constructionist characterisation of WFs and MMs in English, and for their treatment as instance members of a Sceptical Small construction.

The theoretical arguments put forward in this section will be backed up by the data-based analyses reported in section 3.

3 Empirical analysis of *why*-fragments and Mad Magazine sentences

This section reports on two empirical case studies involving, respectively, WFs and MMs in contemporary English, that aim to back up the theoretical arguments outlined in section 2. The investigation on WFs in section 3.1 uses a multivariate statistical methodology in an attempt to determine whether or not the construction-specific sceptical interpretation is favoured by the construction's formal features. Section 3.2 deals with MMs and the evolution of their formal and meaning-related characteristics in contemporary English.

3.1 *Case study 1:* Why-*fragments*

The aim of this first case study is to explore the diachronic evolution of the specific interpretation of WFs in terms of their potential constructionalisation as a special form-meaning pair distinguishable from their corresponding complete *why*-questions. The methodology consisted of a fixed-effects regression analysis to determine the impact of a series of linguistic factors on the semantics of WFs, and to assess whether the pairing of form and meaning in the construction is subject to variation in recent times. The analysis seeks to answer the following research questions:

1. Is the construction-specific interpretation of WFs motivated by linguistic factors?
2. Is the specific interpretation favoured in contemporary English?

3.1.1 Methodology

The data for the study were retrieved from samples of the two editions of the *British National Corpus*: BNC1994 (BNC Consortium 2007) and BNC2014 (Love *et al.* 2017; Brezina *et al.* 2021). These two corpora contain approximately one hundred million words of written and spoken British English from the periods 1960s–93 and 2012–16, respectively. Initially, both the spoken and written components of the two corpora were surveyed; however, the latter was discarded because the only platform where Written BNC2014 was available at the time of the study, #LancsBox X 1.1.0 (Brezina & Platt 2022), did not permit a search for punctuation and the contextual size was too limited to allow a qualitative analysis of the data. Additionally, the spoken components of BNC1994 and BNC2014 (henceforth, Spoken BNC2014) were found not to be comparable, as the former comprises both 'formal encounters from particular institutional settings' and spontaneous conversations, while the latter consists of informal conversations only (Love *et al.* 2017: 321). It was thus necessary to restrict the analysis to the 'conversational' part of spoken BNC1994, what is known as the 'demographically sampled' component (henceforth, BNC1994 DS). The data for the case study were thus retrieved from two samples of different sizes: BNC1994 DS, containing 4,233,962 words and 610,557 so-called 's-units' (i.e. elements 'more or less

equivalent to sentences'; Burnard 2007); and Spoken BNC2014, containing 11,422,617 words and 1,197,091 utterances (BNC 2014 *User manual* 2018: 17).[6]

The data were collected using the CQPweb interface[7] and four queries summarised by the following pattern:

why + (+) (+) (+) \?

The search string is designed to look for the interrogative proform *why*, followed by a minimum of one and a maximum of four words before a question mark. A threshold of four words was set in view of the negligible occurrence of utterances containing remnants of five or more words. The parentheses signal optional elements, and the backslash is used to 'escape' the use of the question mark as a wildcard and match the literal character.[8] The limitations of the transcription of the Spoken BNC2014 data, in which '[t]he only feature of written punctuation retained … is the question mark … to mark questions' (BNC 2014 *User manual* 2018: 37), restricted the scope of the study to interrogative fragments only.

A total of 1,034 hits were extracted from BNC1994 DS and 2,676 from Spoken BNC2014. Manual pruning of the data was necessary to discard instances (in bold below) that do not qualify as WFs, such as non-fragmentary sentences or clauses (30)–(31), fragments with a *wh-* (32) or a finite (33) remnant, and fragments which constitute two (potentially) different questions (34)–(35) or turns by different speakers (36):

(30) [– I mean just putting it out there]
    – **why would you do that?** (Spoken BNC2014: 256 S7GJ)
(31) I wonder why not, wonder **why that is?** (BNC1994 DS: 97 KD0)
(32) [– why didn't we go in for it?]
    – **why what?**
    [– why didn't we do a sheep?] (Spoken BNC2014: 453 ST64)
(33) [– she was going to try and throw it out of the window but she didn't have the strength.]
    – What she threw, **why would?** (BNC1994 DS: 85 KC9)
(34) [– Do you know anybody wanting baby rabbits?]
    – No, **why you've got some?** (BNC1994 DS: KC2 3376)
(35) [– Oh you shouldn't do it right up at the top.]
    – **Why? American?** (BNC1994 DS: 43 KCD)
(36) [– well damp comes through that wall it's **why the**]
    – >>**which wall?** (Spoken BNC2014: 163 S7NJ)[9]

---

[6] Information retrieved from www.natcorp.ox.ac.uk/docs/URG/BNCdes.html#BNCcompo
[7] Information retrieved from https://cqpweb.lancs.ac.uk
[8] See https://cqpweb.lancs.ac.uk/doc/cqpweb-simple-syntax-help.pdf
[9] The symbol >> indicates that 'the beginning of the utterance overlaps in time with the end of the previous overlap' (BNC 2014 *User Manual manual* 2018: 50).

Non-prototypical WFs without explicit remnants (37) or followed by either the negator *not* alone (38) and/or interpersonal/textual (non-ideational)[10] material (39)–(42) were also excluded:

(37) [– Yeah, I dunno why]
     – **why?**
     [– I dunno, just think she will] (BNC1994 DS: KD8 6924)

(38) [– I wouldn't want it]
     – **why not?** (Spoken BNC2014: 66 S59W)

(39) [– Don't want these black-currant sweeties.]
     – **Why babe?** (BNC1994 DS: 3 KB6)

(40) you know why waste them? Why **why you know?**(Spoken BNC2014: SJNB 2222)

(41) [– >>cos you can't]
     – **why mother?** (Spoken BNC2014: SKRC 421)

(42) [– I don't make a habit of admitting to it.]
     – **Why ever not?** (BNC1994 DS: KD8 10397)

Unfinished utterances (43) and unclear instances (44) were excluded as well, together with repetitions of the same fragment by the same speaker. In the latter case, only the first occurrence of the fragment was considered, as illustrated in bold in (45).

(43) [– I mean there's no reason]
     – **why d-?** why is four the critical number then? (Spoken BNC2014: 249 SF3V)

(44) [– I've got a bit of spare time I can do it]
     – **why Scotland --UNCLEARWORD?** (Spoken BNC2014: 365 SUVS)

(45) – **why the long pause?**
     [– >>--UNCLEARWORD I'm not hungry]
     [– >>when when did you have your breakfast?]
     – >>why the long pause?
     [right I'll have to go –UNCLEARWORD]
     – >> why the long pause? (Spoken BNC2014: SCP7 679)

After this manual pruning of the data, a total of 82 valid instances were identified in BNC1994 DS and 177 in Spoken BNC2014.

3.1.2 *Variables analysed*

Variation in meaning of WFs was modelled by running a fixed-effects regression, which controlled for the impact on the semantic nuances expressed by the fragments (dependent

---

[10] The notions of 'interpersonal' and 'textual' elements are borrowed from Systemic Functional Linguistics (see, for example, Halliday & Matthiessen 2014: §5.1.3). Participants (subjects, objects, prepositional/adverbial complements and predicatives) and circumstances (adverbials) are linked to the so-called 'ideational' metafunction of the grammatical system. Conjuncts and conjunctions are classed as textual expressions, while interpersonal elements include disjuncts, auxiliary-centric phrases conveying modality, aspect and polarity information, and semantically bleached comment clauses and vocatives, among others (Halliday & Matthiessen 2014: §3.4).

variable (i)) of a series of relevant variables surveyed in previous literature (independent variables (ii)–(iv)).[11]

(i) 'semantics' codes the type of nuance expressed by the WF. As pointed out in section 2, all fragments have a 'canonical' interpretation, i.e. a propositional meaning which is equivalent to that of the non-reduced *why*-interrogative construction, as illustrated in (46). However, some WFs may also convey a so-called 'construction-specific' nuance, which cannot be directly derived from the complete *why*-sentence or the reduced construction. This pragmatically enriched interpretation usually involves a 'modal' nuance (47)–(48), which resembles the meaning of *should* or *would*, and 'uniqueness' (49)–(50). Both imply a certain scepticism with respect to the previous proposition and are, therefore, included under the subcategory 'specific'. All the (contextualised) WFs in our database underwent evaluation by a minimum of two native speakers. They assessed the equivalence of these forms to the following alternatives: orthodox questions (e.g. 'why doesn't everybody get better rewards?'), versions involving modal *should* ('why shouldn't everybody get better rewards?'), versions including specifically ('why that way specifically?') and versions completed by 'Doesn't everybody get better results? I can't believe it!'). The successful inter-rater results enabled us to reliably classify 30.89 per cent (80 tokens) of the WFs as conveying the sceptical interpretation.

(46) [– Hurry up with the scissors Pauly.]
        – **Why the rush?** (BNC1994 DS: KD0 499)
(47) [– you get better rewards for being rubbish at your job at that place]
        [– well for her yes but]
        – well **why not for everybody then?** (∼ 'why shouldn't everybody get better rewards?') (Spoken BNC2014: S457 592)
(48) [– That's stupid! Why make any comment whatsoever?]
        – Yeah. **Why not just keep quiet?** (∼ 'why shouldn't they just keep quiet?') (BNC1994 DS: KBD 4944)
(49) [– round two's coming on one week's time]
        – **why not now?** (∼ 'one week's time vs now') (Spoken BNC2014: SUWR 58)
(50) [– Knives and forks should go in that way. Right?]
        – **Why that way?** (∼ 'that way vs any other way') (BNC1994 DS: KCL 703)

(ii) 'corpus', with values BNC1994 DS and Spoken BNC2014, indicates the period the instances analysed belong to and allows for a diachronic analysis of the data. 68.34 per cent (177 instances) of the data come from Spoken BNC2014.

(iii) 'category (of the remnant)' controls for the type of phrase that follows the *why*-proform. Three broad categories were identified: '(pro)nominal' constituents, as in

---

[11]   In prior stages of this research, a fourth independent variable was considered: 'island violation' (see section 2 for further information). However, this factor proved not to be statistically significant and, hence, was not included in the statistical modelling of the data (see Fernández-Pena & Pérez-Guerra 2024).

(51)–(52), which account for 38.61 per cent of the data (100 tokens); nonfinite clauses (53), accounting for 45.56 per cent of the database (118 instances); and 'other' phrases (54), a subcategory representing the least frequent types of remnants (i.e. adverbial, adjectival and prepositional phrases, as well as Small Clauses), which covers the remaining 15.83 per cent of the results (41 tokens).

(51) [– it's just ten hours of rain or like ten hours of forest sounds ten hours of like it's usually]
        – >>**why ten hours?** (Spoken BNC2014: SY7T 167)
(52) [– Sarah [gap:name], Claire [gap:name] what do you reckon to her?]
        – **Why her?** (BNC1994 DS: KSV 3328)
(53) if you're going to be leaving at five **why stay?** (Spoken BNC2014: SLBD 711)
(54) when did Jesus in fact become Christ? And **why then?** (BNC1994 DS: KBX 1594)

(iv) 'category mismatch' refers to instances containing a discourse-new remnant, i.e. a remnant whose category differs from that of the antecedent. For example, 'category mismatch = yes', in (55), indicates that the antecedent of the adverbial phrase *that early* is the prepositional phrase *in January*. Category mismatch occurs in only 2.32 per cent (6 tokens) of the data retrieved from both corpora. Cases in which there is no mismatch (i.e. 'category mismatch = no'), such as (56), represent 61 per cent (158 instances) of the database, while informatively new remnants, such as (57) (i.e. 'category mismatch= NA') account for slightly more than one third of the examples retrieved (95: 36.68%).

(55) why are you advertising in January for September? […] **why that early?** (Spoken BNC2014: S5JX 523)
(56) [– I've just been offered on Tinder to have your breakfast made half naked][…]
        – **why half naked?** (Spoken BNC2014: SMC2 525)
(57) He's been like that all his life, **why bother now?** (BNC1994 DS: 255 KPX)

3.1.3 *Discussion*

As mentioned in section 3.1.1, the number of instances retrieved is higher in Spoken BNC2014 than in BNC1994 DS. The normalised frequencies (per million sentences)[12] shown in table 1 confirm that there is an increase in the frequency of WFs from the period 1960s–93 (nf/pms: 134.3) to the 2010s (nf/pms:147.9). This trend, however, was found to be not yet statistically significant ($\chi^2(1)=0.43$, $p=.512$). The results from the regression model reported below offer additional insight into this incipient increase.

To gauge the impact on the meaning of WFs of the independent variables identified above ('corpus' included), a fixed-effects binomial regression model was fitted.[13] After confirming that (multi)collinearity of the variables analysed was not potentially

---

[12] Normalisation of the data was carried out using the total number of so-called 's-units' in the case of BNC1994 DS (i.e. 610,557; Burnard 2007) and 'utterances' in that of Spoken BNC2014 (i.e. 1,197,091; see BNC 2014 *User manual* 2018: 17).
[13] The regression model used the statistical software R (R Core Development Team 2022) and the functions glm() from the 'stats' package (R Core Development Team 2022) and 'lrm' from the 'rms' package (Harrell Jr 2022).

Table 1. *Raw and normalised frequencies of WFs*

BNC1994 DS	Spoken BNC2014
82	177
nf/pms: 134.3	nf/pms: 147.9

harmful (i.e. 1.089–1.248),[14] backward stepwise elimination[15] was run in order to obtain the model that best accounts for the variation in the data. None of the variables was discarded as non-significant, which indicates that the three predicting variables considered – 'corpus', 'category (of the remnant)' and 'category mismatch' – sufficiently explain variation in the semantics of WFs. The evaluation of the optimal model confirmed acceptable discrimination and classification accuracy (C-index = 0.721) and explanatory power (Nagelkerke R^2 = 0.151). Table 2 reports the output of the model, with predicted odds for specific meaning.

The output of the model shows that, with all predictors at their baseline levels, there is an overall, albeit statistically non-significant, preference for specific meanings (p=.8303). The estimates for the fixed effects indicate that this preference for construction-specific interpretations is significantly favoured by the most recent data ('corpus:bnc2014'), nonfinite remnants ('category:nonfinite') and discourse-new remnants ('category: mismatch:yes'), i.e. cases in which the category of the remnant differs from that of the antecedent.

Figures 2 to 4 illustrate the trends attested by the model by plotting the predicted probabilities for the semantics of the WFs in each case: the higher the probability of the specific interpretation, the closer to the topmost part of the graph; a higher incidence of canonical interpretations, on the other hand, is reflected in the bottom part of the graph.

Figure 2 corroborates the stronger association of nonfinite WFs and specific nuances, as well as the marked preference of 'nominal' and especially 'other' WFs for canonical interpretations.

Figure 3 illustrates the impact of the corpus on the probability of specific (vs canonical) meanings. The graph confirms an increase in frequency in relation to specific interpretations from BNC1994 to BNC2014: the WFs in Spoken BNC2014 are more strongly associated with sceptical nuances than the data in BNC1994 DS (p=.0355).

Figure 4 plots the predicted probabilities of the (mis)match in category on the interpretation of WFs. The results confirm a significant correlation between new remnant categories and specific nuances (p=.0100) and the consequent preference for canonical nuances when the remnant and the antecedent share the same category.

---

[14] (Multi)collinearity was controlled for using the function 'vif' from the 'car' package (Fox & Weisberg 2019).
[15] Backward stepwise elimination was carried out using the function 'step' from the 'MASS' package (Venables & Ripley 2002).

Table 2. *Output of the fixed-effects binomial regression model (predicted odds are for specific meaning)*

Fixed-effects binomial model
Response variable: semantics (canonical|specific) (179:80)

Overall model statistics
Classification accuracy: *C*-index: 0.721
Explanatory power: Nagelkerke R^2: 0.151

Coefficients	Estimate	Std error	z value	p-value	
(Intercept)	0.1977	0.9222	0.214	0.8303	
corpus:bnc2014	1.0747	0.5111	2.102	0.0355	*
category:nonfinite	0.9338	0.4158	2.245	0.0247	*
category:other	-1.4700	0.7676	-1.915	0.0555	.
category:mismatch:yes	2.4288	0.9433	2.575	0.0100	*

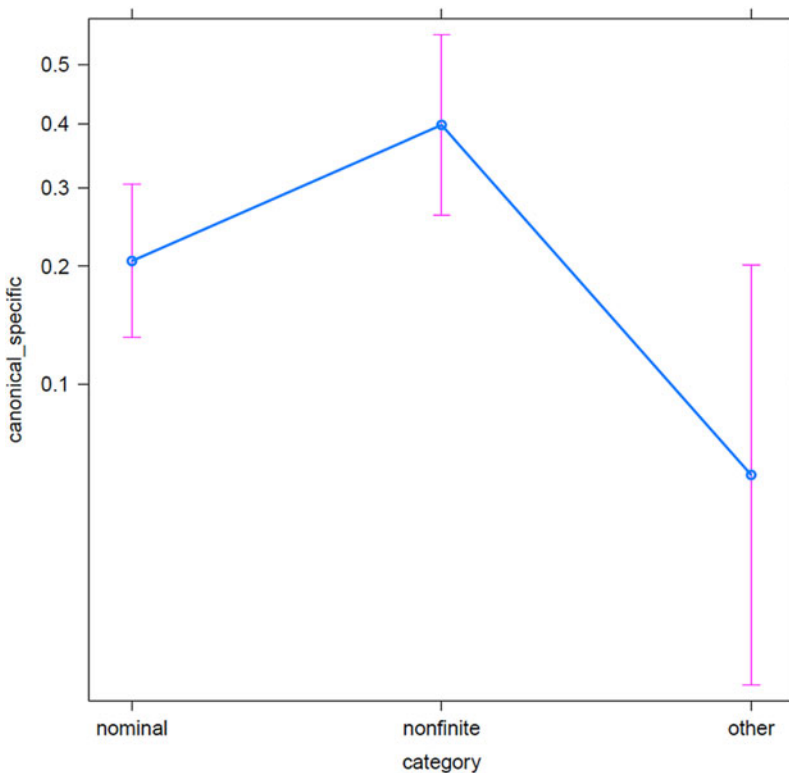Significance codes: 0 '***'; 0.001 '**'; 0.01 '*'; 0.05 '.'; 0.1 ' '.

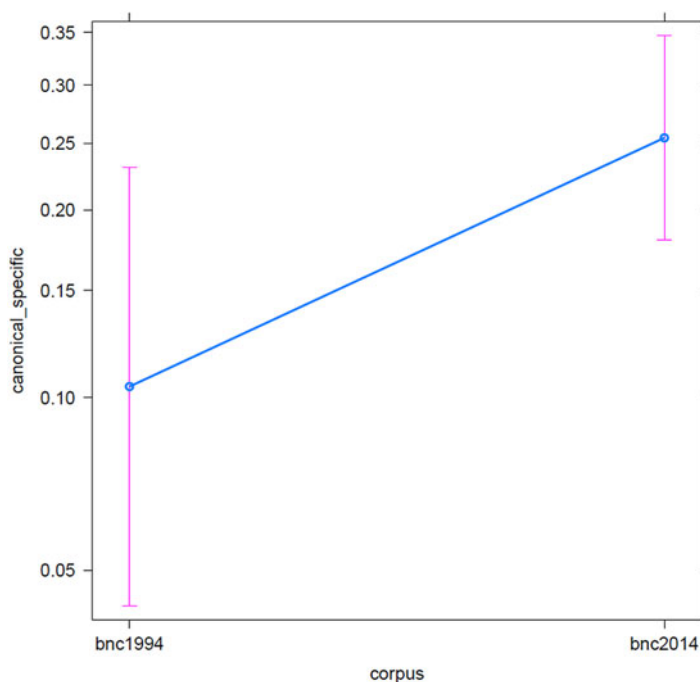Figure 2. Effect of 'category' on the predicted probability of specific meanings

Figure 3. Effect of 'corpus' on the predicted probability of specific meanings

Taken together, the results of figures 3 and 4 and the (timid) 'diachronic' increase in frequency illustrated in table 1 seem to be pointing to an ongoing trend: not only do we observe a change in that it is the novel remnant categories, the ones whose form differs from that of the antecedent, that tend to denote construction-specific nuances, but the data also evince a semantic change in recent times, with specific interpretations having significantly increased in frequency from the late twentieth century to the early twenty-first century. These results would seem to suggest an incipient process of constructionalisation in terms of fixation in the pairing of discourse-new remnants with specific nuances, and an increase in the frequency of specific interpretations from the late twentieth century to the early twenty-first century.

3.2   Case study 2: Mad Magazine sentences

The aim of the second case study is to explore MM constructions in contemporary English, specifically their potential constructionalisation as special form-meaning pairings in recent times. In this case, for reasons explained in section 3.2.1, a more qualitative exploratory methodology has been used. The research question that this analysis seeks to answer revolves around the existence of evidence of a constructionalisation process of MMs in contemporary English in relation to form and meaning. To answer the first part of the question, category, syntactic function and
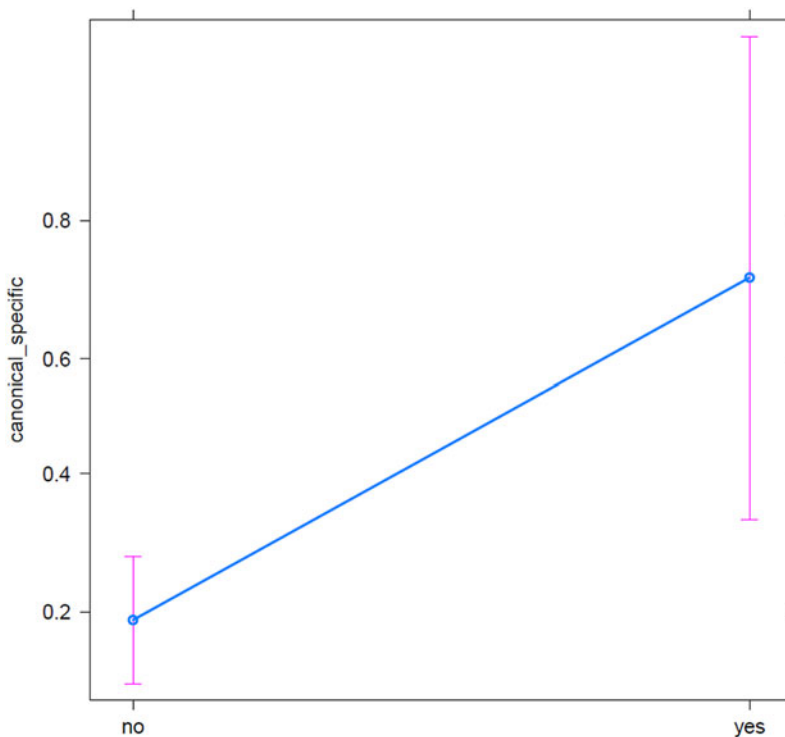
Figure 4. Effect of 'category mismatch' on the predicted probability of specific meanings

textual (i.e. register) distribution of the construction were assessed. To answer the second, the analysis examined polarity, referential status, subjective reference and potential sceptical interpretation.

3.2.1 Methodology

To ensure comparability of the results, this case study was initially intended to explore MMs in samples of the BNC1994 DS and Spoken BNC2014. However, owing both to the impossibility of searching for punctuation using #LancsBox X 1.1.0 (Brezina & Platt 2022) at the time the research was carried out and to the low frequency of instances obtained in BNC1994 DS and Spoken BNC2014 using CQPweb,[16] the data for the study were ultimately retrieved from the *Corpus of Contemporary American English* (COCA, 1990–2019; Davies 2008–). To explore potential diachronic trends, the corpus was divided into two subperiods: 1990–2004 and 2005–19.

As described in section 2, MM constructions consist of an initial element, here labelled 'MM1', which may be an accusative pronoun or a name, and a second element, 'MM2', which is usually a nonfinite verbal phrase (e.g. *me, worry?*), an adjectival phrase (e.g. *me*

---

[16] 255 (unfiltered) instances were obtained from BNC1994 DS but only 13 from Spoken BNC2014.

worried?), a prepositional phrase (e.g. *You, with MMs?*) or a nominal phrase (e.g. *me, beer?*). These two parts of MMs may occur in reverse order as well (i.e. MM2 MM1, as in *a doctor, your son?*) and may be separated by a comma or a negator (e.g. *You(,) (not) hungry?*). To avoid as much noise in the data as possible, the search focused on MMs with pronominal subjects which followed the most frequent order MM1-MM2 (see Lambrecht 1990: 221 on MM1/MM2 reversibility in examples such as *Wear a tuxedo, him?*), and were filtered by the presence of initial and end-position punctuation. Owing to technical problems associated with query length,[17] prepositional MMs could not be retrieved. The following three patterns summarise the 60 different queries used to retrieve the examples:

i. PUNC PRON$_{ACC}$ (,) (NEG) ADJ ?|!|.
ii. PUNC PRON$_{ACC}$ (,) (NEG) VERB ?|!|.[18]
iii. PUNC PRON$_{ACC}$ (,) (NEG) (DET|POSS) NOUN ?|!|.

These three patterns yielded instances of (i) adjectival (58), (ii) verbal (59) and (iii) nominal (60) MMs, respectively. The items in parentheses represent optional elements and those separated by a vertical slash correspond to the three different end punctuation marks by means of which MMs were filtered.

(58) [– I didn't want to worry you.]
        –**Me, worried?** (COCA:1997:TV Nash Bridges)
(59) [– Archie, do you dance?]
        – Oh, my, **me, dance?** No, no. (COCA:2017:MOV …e & Her Sisters in a Puppy Chase)
(60) [– The guy's one large muscle, and we're talking head to toe.]
        – **Me, all muscle?** (COCA:1997:TV Duckman: Private Dic…)

The total number of hits obtained was 130,206: 4,499 adjectival, 108,568 verbal and 17,139 nominal. In view of the high number of tokens, the analysis was restricted to samples of 500 random examples in patterns with a token frequency in excess of 500. The total number of instances manually analysed thus amounted to 5,394: 1,381 adjectival, 1,678 verbal and 2,335 nominal. Manual pruning of the data discarded irrelevant instances, such as: (61), where the pronoun and the noun phrase do not form an MM construction; (62), where *her* does not function as an accusative pronoun but as a possessive determiner; and (63), which contains a finite form.

---

[17] The queries for prepositional MMs involved highly frequent items and, in these cases, the search engine limits the string to five words, making it impossible to retrieve the data.

[18] Although verbal MMs are nonfinite by definition, the decision was taken to search for all types of verbal forms after detecting that the verbal forms of prototypical MMs, such as *What, him worry?* (COCA:2012:SPOK CBS News Sunday Morning) were tagged as VV0, i.e. as non-third person singular verb forms. Similarly, the search returned a high number of examples of verbal MMs where the verbs received the POS-tag corresponding to singular nouns (i.e. NN1), as in *What, me worry?* (COCA:1992:NEWS SanFranChron) or *Me lie?* (COCA:2011:TV Burn Notice). Given their statistical significance, these tokens were included in the analysis of the data.

(61) [– Why should I do anything?]
         – They're going to kill me, **you, either way**. (COCA:2008:MOV Iron Man)
(62) [– Her husband Jonas had been put under a temporary ban[.]]
         – **Her own husband?** (COCA:2002:FIC NewEnglandRev)
(63) [– 'I knew that.']
         – '**You knew?**' # He looked a little defensive. (COCA:2012:WEB craphound.com)

Cases of parenthetical structures (64) and vocatives (65) were also excluded:

(64) but everyone swears it was getting larger, **you know!** (COCA:1999:MOV The Alien Legacy)
(65) [– Let him loose.]
         – **You lunatic!** (COCA:1991:TV MacGyver)

After manual pruning, the total number of remaining valid tokens was 152: 75 adjectival, 48 verbal and 29 nominal. Since the definitive database was too small to permit a statistical multivariate analysis, a more frequency-related qualitative exploration of MMs conveying the specific interpretation of scepticism, surprise, disbelief and/or scorn explained in section 2 was used instead.

3.2.2 *Variables analysed*

Based on previous research on MMs, the following variables were considered in this study:

(i) 'subjective reference of MM1' refers to the type of accusative subject in the MM sentence. In our data, most of the pronominal MM1s are either first person singular (57; 37.50%) (66) or second person (90; 59.21%) (67), and much less frequently first person plural (3; 1.97%) (68) and third person singular (2; 1.32%) (69):

(66) [– Have you been crying In the restroom?]
         – **Me, crying??** (COCA:2006:MOV …ve 2006 Academy Awards Post Show)
(67) [– For a mere thirty francs apiece.]
         – **You laugh?** I can get fifty for them in Paris. (COCA:2000:MOV Chocolat)
(68) [– We should stay here all week. What if we got bored?]
         – **Us, bored?** (COCA:2018:MOV Irreplaceable You)
(69) [– She's worried, the way anyone is when they've only got one chance to get something right. Worried, but not afraid.]
         – **Her, afraid?** Her, sunglancer, moondancer, heaven gate crasher, astromancer? (COCA:2019:FIC Analog)

(ii) 'category of MM2' codes the type of element that accompanies the accusative pronoun: adjectival (70), nominal (71) or verbal (72)–(73). Almost half of the instances retrieved are adjectival (75 tokens; 49.34%), one third (48; 31.58%) are verbal and the remaining 19.08 per cent (29 tokens) are nominal.

(70) [– I mean, I'm sorry for the man, but I ain't got nothing to say.]
         – **You sorry?** You sorry for him? (COCA:2002:TV The Wire)

(71) [– And you are my son.]
     – **You my father?** I'm your son? I didn't know. (COCA:2010:MOV The Last Godfather)
(72) [– What do you think about adopting?]
     – **Me, adopting?** (COCA:2005:MOV Say Uncle)
(73) [– Are you badly hurt?]
     – **Me, hurt?** I'm tired, that's all. (COCA:1997:MOV … Walt Disney's Sleeping Beauty)

(iii) 'function of MM2' categorises the syntactic function of the second member of the MM construction. MM2 may fulfil three different functions: predicative, which accounts for 61.84 per cent of the database (94 tokens), as in (74); argument (either object or prepositional complement) (75), accounting for 5.26 per cent (8 instances); and adverbial (76), of which only 2 instances (1.31%) were found.[19]

(74) [– Are you sure you're up for it? – Yeah.]
     – **You sure?** (COCA:2011:TV Switched at Birth)
(75) [– I'm gonna round it down to an even 8 grand.]
     – **It, $8,000?** (COCA:2008:TV Unhitched)
(76) [– Well, say you'll come.]
     – **Me this weekend**. (COCA:2009:TV Brothers & Sisters)

(iv) 'referential status of MM2' controls for textually referring instances or phoric referents (77) and non-referential or exophoric referents (78). The vast majority of MMs were found to be referential (122; 80.26%).

(77) [– If only you could hunt. I know how to hunt.]
     – **You, hunt?** (COCA:2011:MOV Oka!)
(78) [– My friend said that Davos is like an Outward Bound for the rich and famous: they're forced to come here without their entourages because there are so few hotel rooms.]
     – What? **Me, sleep?** (COCA:1999:MAG Inc.)

(v) 'polarity' codes the potential impact of the presence (79) and absence (80) of a negator before MM2. Only a small minority of the MMs retrieved show negative polarity (5; 3.29%).

(79) [– You'll be doing that the rest of your life, said Chicken, unless you decide to break the rules.]
     – # What! exclaimed Rooster, **me not crow?** (COCA:2001:FIC LitCavalcade)
(80) [– She said you sounded crazy. She didn't want to face you.]
     – **Me, crazy?** (COCA:1995:MOV The Brothers McMullen)

(vi) 'register' controls for the text types represented in COCA in which the MMs occur: i.e. 'blog', 'web', 'TV/Movies', 'spoken', 'fiction', 'magazine', 'newspaper' and 'academic'. To simplify these categories, owing to their speech-based nature, 'blog',

---

[19] Akmajian (1984: 13) observes that MMs are incompatible with time adjuncts. By contrast, Osawa (2011) contends that MMs may contain time adverbials and even uses this fact to support the need for a null Tense functional category in MMs. Although our data do not contain examples with adjuncts co-occurring with other MM2 arguments, the results include two valid instances of MM2 time adjuncts, as in (76).

'web', 'TV/Movies', 'spoken', 'fiction' and 'magazine' have been grouped into a single subcategory termed 'speech-related registers',[20] where almost all of the MMs occur (150; 98.68%). Only two instances of MMs were found in 'newspaper'.

3.2.3 Discussion

One of the main aims of this case study is to identify evidence of a constructionalisation process of MMs in contemporary English in relation to form by assessing their distribution in terms of category, syntactic function and register. Figures 5 and 6 illustrate the distribution of, respectively, the different categories and functions of MM2 in the two subperiods of COCA. Figure 5 shows that adjectival MMs are, unsurprisingly, the most frequent category in both periods, followed by verbal and nominal. The minimal differences observed between the late twentieth century and early twenty-first century are not statistically significant ($\chi^2(2)$=.302, p=.859866).

The results in figure 6 show that MM2 mostly functions as predicative, argument and oblique MM2s occurring significantly less frequently. The most important finding, however, is that the patterns attested in both subperiods are almost identical; the diachronic analysis is, therefore, statistically non-significant (predicative versus other; Fisher exact test: p=.7417).

The third variable explored in relation to the potential effect of the constructionalisation of MMs on their form is their distribution across the different registers represented in COCA. As shown in figure 7, only two instances of MMs are attested in a non-speech-related genre in the late twentieth century. As with the previous variables, MMs pattern alike in both subperiods.

The second dimension of this case study concerns the impact of the potential constructionalisation of the MMs on their meaning. To answer this question, four different factors were explored: MM polarity, MM referential status, subjective reference of MM1, and sceptical interpretation of MMs. Starting with polarity, the results show that the vast majority of MMs in both subperiods of COCA have positive polarity (98.60% in 1990–2004 and 94.90% in 2005–19), with trends remaining stable over time (Fisher exact test: p=.3675), as shown in figure 8.

Regarding the referential status of MM2, the data show, as expected (Lambrecht 1990), that most MM2s are referential both in the late twentieth century and the early twenty-first century. Nonetheless, in our data exophoric MMs, which, as pointed out by Donaldson (2013: 17), are possible, are not as rare as might be expected (16.22% in 1990–2004 and 23.08% in 2005–19) – see figure 9. Once again, no significant diachronic changes were attested ($\chi^2(1)$=1.1283, p=.288137).

---

[20] The decision to group all speech-related registers into a single subcategory was based on the corpus description of the subregisters: 'blogs' 'represent a subset of the texts from the United States in the GloWbE corpus'; 'TV/ Movies' are 'as informal (or more informal) than actual spoken data', 'spoken' represents 'unscripted conversation from … TV and radio programs'; 'fiction' is composed of '[s]hort stories and plays from literary magazines, children's magazines, popular magazines, … books …, and movie scripts'; and 'magazine' contains magazines from 'a good mix … between specific domains (news, health, home and gardening, women, financial, religion, sports, etc.)'. For further information, see www.english-corpora.org/coca/
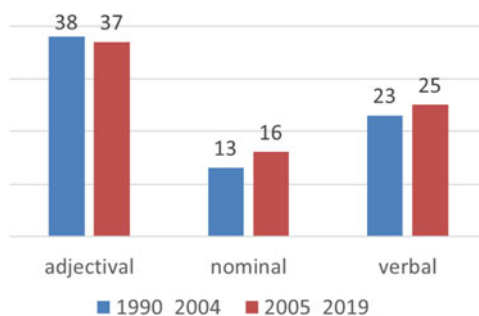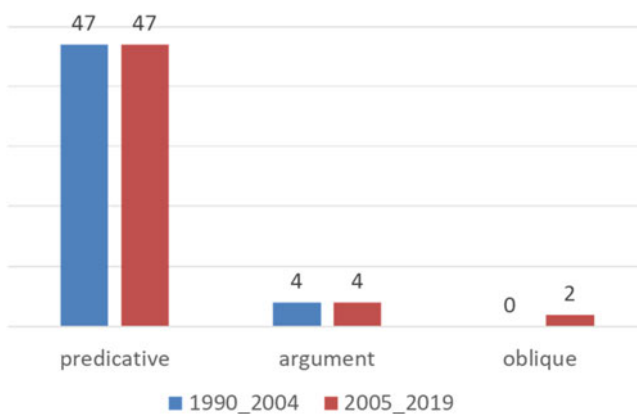
Figure 5. Distribution of MM2 category

Figure 6. Distribution of MM2 function

Regarding the subjective reference of MM1s, the data reveal that most MMs take either first person singular or second person pronouns as subjects. Figure 10 also shows a slight change between the late twentieth century and the early twenty-first century, with *me* occurring as the preferred option in the former (52.70%) and *you* in the latter (73.08%).

To sum up, this case study did not reveal significant diachronic differences in the form and meaning of MMs in contemporary American English. In terms of form, the results show that MMs most frequently contain adjectival and predicative MM2s, and that, in general, MMs occur in speech-related registers. As regards meaning, the vast majority of MMs proved to have positive polarity, referential status of MM2 and first-/second-person pronominal MM1s. On the whole, unlike WFs, the evidence obtained from COCA seems to indicate that MMs in contemporary American English are already constructionalised. The manual analysis of the whole sample of MMs supports this claim, as it confirms that all of them convey not only the meaning of a canonical question but, more importantly, scepticism about a previous proposition.
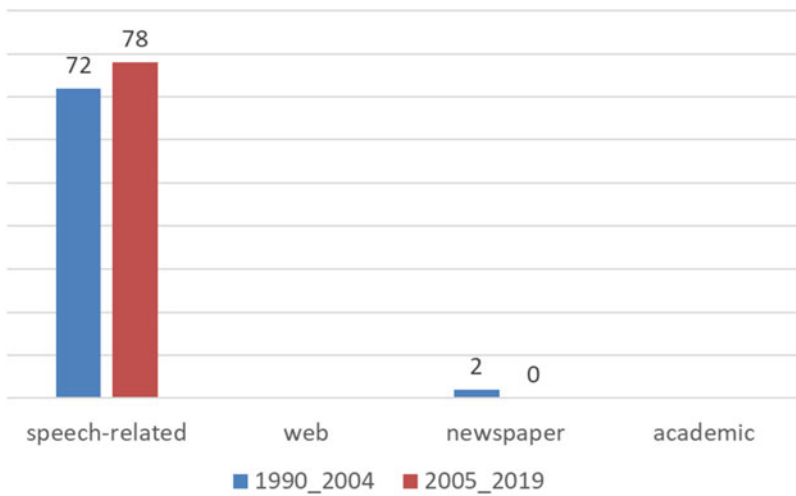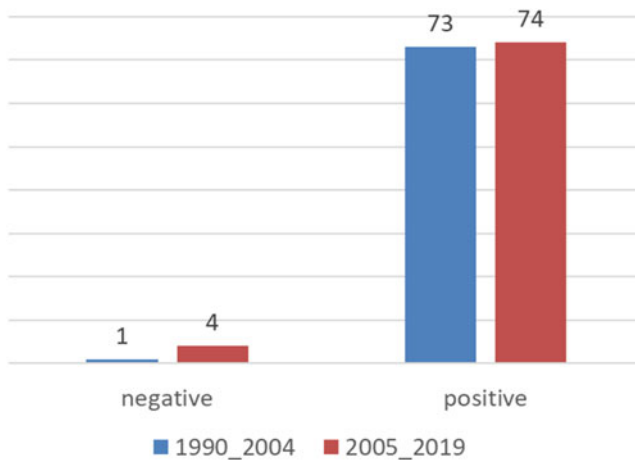
Figure 7. Distribution of MMs by register

Figure 8. Distribution of MM polarity

4 A constructional account of *why*-fragments and Mad Magazine constructions: the 'sceptical small' construction

Construction Grammar is an appropriate framework for the analysis of fragments. Fragments can 'become entrenched in a speaker's mental grammar provided they occur with sufficient token frequency' (Bauer & Hoffmann 2020: 241), in keeping with the constructionalist premise that our knowledge of language is based on language use (Goldberg 2006; Bybee 2013). The usage-based perspective assumes, among other principles, that speakers are 'at once impressively creative and impressively repetitive' as far as their linguistic productions are concerned (Goldberg 2013: 26), and that these
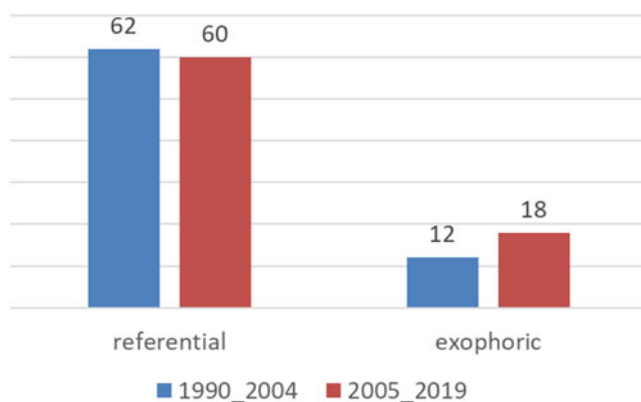
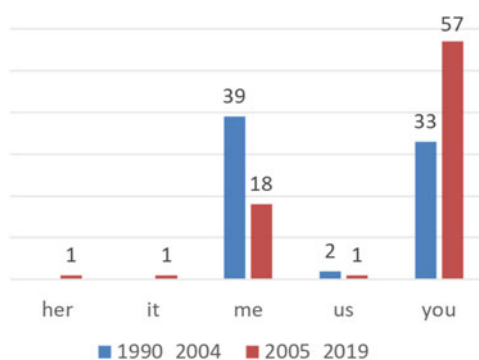Figure 9. Distribution of referential and exophoric MMs

Figure 10. Distribution of the different types of pronouns in MM1

processes of creativity and repetition go beyond the level of individual lexical items to affect entire constructions (see Van de Velde *et al.* 2015). The analysis proposed here adopts the definition of 'construction' offered by so-called Cognitive Construction Grammar:

> Any linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognizable to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency. (Goldberg 2006: 5)

This study has shown that formal (syntactic design, constituents' categorial status, textual distribution) and meaning (polarity, referentiality, canonical versus sceptical interpretation) features are conventionalised in WFs and MMs in a non-compositional way, and that these expressions may therefore be classed as special pairings of form and meaning, that is, constructions in their own right, distinct from complete questions.
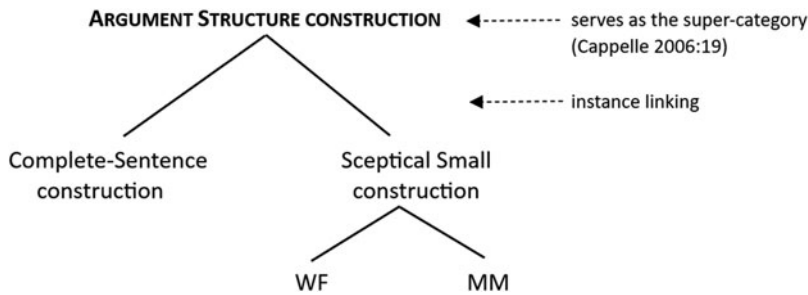
Figure 11. WFs and MMs as examples of Argument Structure Constructions

Firstly, since WF and MM fragments are reduced or 'small' expressions that are functionally (i.e. semantically, pragmatically, communicatively) equivalent to complete questions in discourse, competing with complete questions when these fragmentary patterns convey what we have called a canonical interrogative interpretation (see section 1), WFs and MMs can be taken as (not necessarily immediate) instances of a super-category (Cappelle 2006: 19) construction, together with the Complete-Sentence/Question construction, as reflected in figure 11.

Secondly, given the formal and semantic/pragmatic similarities between WFs and MMs (i.e. interrogative(/exclamative) intonation, subject optionality, accusative case-marking on subjects, nonfinite predicates, questioning of the proposition conveyed by the action, event or state leading to a specific interpretation of scepticism), this article argues that WFs and MMs may be analysed as instantiations of the same umbrella construction, the so-called Sceptical Small construction. The form-meaning pairing of the two Sceptical Small constructions is summarised in figures 12 and 13.

In both cases, the formal design of these constructions consists of two members (plus introductory *why* in WFs): the first, in the role of subject,[21] is optional; the second is a predicative, argument (object) or oblique constituent (see section 2 for the main syntactic roles of WF remnants and MM2s) or a nonfinite predicate. With regard to the first member, subjects may be either explicit or implicit (unpronounced), and in the former case either referring/phoric or exophoric (or 'deprofiled'; Goldberg 2000). With regard to the second member, the null or nonfinite (explicit) predicate justifies the categorisation of the construction as 'small'. Also, as far as form is concerned, in prototypical WFs and MMs the intonation is interrogative (see section 2). As regards meaning, WFs and MMs may inherit the construction-specific sceptical interpretation of the umbrella Sceptical Small construction in which the subject, if overt, is the Theme, and the remnants in WFs and the MM2s are prototypically attributes or events. Finally, as already mentioned, subjects ('A') in WFs and MMs prototypically convey given, referring, 'EVOKED' information.

---

[21] (Non-subject) Object A_2 members, as in [*Mary kissed Bill and John.*] *Why John?*, are not prototypical instances of WFs.

```
FORM:       PHONOLOGY:     /why₁ (A₂) B₃/₄
                           INTONATION: interrogative(, exclamative)
            MORPHOSYNTAX:           ⎡ PREDICATIVE₃      ⎤
                           why₁ (SBJ₂) ⎢ ARGUMENT₃       ⎥
                                       ⎢ OBL₃            ⎥
                                       ⎣ PREDICATE_nonfinite3 ⎦₄
MEANING:    SEMANTICS:     'SCEPTICAL-interpretation (THEME₂,⎡ATTRIBUTE₃⎤)'₄
                                                            ⎣EVENT₃    ⎦
            INFORMATION STRUCTURE: 'EVOKED'₂
```

Figure 12. The WF construction

```
FORM:       PHONOLOGY:     /(A₁) B₂/₃
                           INTONATION: interrogative(, exclamative)
            MORPHOSYNTAX:           ⎡ PREDICATIVE₂      ⎤
                           (SBJ₁) ⎢ ARGUMENT₂       ⎥
                                   ⎢ OBL₂            ⎥
                                   ⎣ PREDICATE_nonfinite2 ⎦₃
MEANING:    SEMANTICS:     'SCEPTICAL-interpretation (THEME₁,⎡ATTRIBUTE₂⎤)'₃
                                                            ⎣EVENT₂    ⎦
            INFORMATION STRUCTURE: 'EVOKED'₁
```

Figure 13. The MM construction

5   Concluding remarks and avenues for further research

This study has examined two fragmentary constructions, WFs and MMs, in contemporary English, both theoretically and by way of an empirical, corpus-based analysis.

Theoretically, WFs and MMs were categorised, firstly, as 'constructions' from a (Cognitive) Construction Grammar perspective and, secondly, as members of the so-called Sceptical Small construction, whose prototypical defining features are, as regards form, interrogative intonation and 'small' or untensed status, and, from the point of view of meaning, conveyance of a non-compositional sceptical interpretation. This constructionalist approach has several advantages. First, it perfectly addresses the so-called 'enriched' interpretation of WFs and MMs by embedding these constructions within an umbrella construction likely to encompass other small or untensed nonsententials that pave the way for an irrealis sceptical nuance. Second, it accounts for formal (intonational, structural, syntactic) similarities between WFs and MMs. Third, it justifies the infeasibility of an ellipsis/deletion explanation of WFs and MMs, and lends support to a specific nonsentential (from other perspectives, 'base-generated') design of the constructions.

For the corpus-based exploration, two case studies were conducted on WFs and MMs, respectively. The first case study focused on the way in which WFs have evolved in recent times in two corpora of British English: BNC1994 and BNC2014. The data revealed a timid increase in frequency of the WF construction, a significant rise in instances conveying the sceptical interpretation across time, and a preference for discourse-new

remnant categories, indicating ongoing constructionalisation in contemporary British English. The second case study dealt with variation in relation to MMs in the periods 1990–2004 and 2005–19, with data retrieved from a corpus of American English (COCA). Focusing on sceptical MMs only, the study explored the evolution of formal and meaning-related features, and concluded that MMs do not exhibit statistically significant differences in relation to their linguistic design, textual distribution, polarity or the referential status of constituents across time. These findings corroborate that the MM construction has already been constructionalised in contemporary American English. In conclusion, this study illustrates that statistical methods may be used to evaluate and compare construction types, linguistic variation and processes of constructionalisation.

Further research will be required (i) to check the entrenchment of the sceptical interpretation in specifically those examples that do not tolerate reconstruction via deletion or ellipsis; (ii) to investigate WFs in American English, MMs in British English and both constructions in other (at least, inner-circle) varieties in order to evaluate the weight of dialectal factors; (iii) to find diachronic support in explorations in the immediately preceding periods, for example, in Late Modern English; (iv) to carry out research on akin linguistic productions in informal settings and in media typically reporting non-standard variations; and (v) to examine and asses other patterns that may fit the criteria for the Sceptical Small construction, such as Siemund's (2018: §10.1.5) 'incredulity response' structures (e.g. *Do this thing – never!*; see footnote 3) and even imperatives, which, as Lambrecht (1990: 215) argues, have optional subjects, are not tensed and 'force an irrealis interpretation of the state, event or action'.

Authors' addresses:

Department of English, French and German
Universidade de Vigo
Campus Universitario s/n
36310 Vigo
Spain
yolanda.fernandez@uvigo.gal
jperez@uvigo.gal

References

Akmajian, Adrian. 1984. Sentence types and the form-function fit. *Natural Language and Linguistic Theory* 2, 1–23.

Alves-Castro, Marian 2020. 'Why not' and other patterns of negative *why*-stripping. In Lourdes López Ropero, Sara Prieto García-Cañedo & José Antonio Sánchez Fajardo (eds.), *Thresholds and ways forward in English studies*, 168–78. Alicante: Publicaciones de la Universidad de Alicante.

Bauer, Eva-Maria & Thomas Hoffmann. 2020. Turns out is not ellipsis? A usage-based construction grammar view on reduced constructions. *Acta Linguistica Hafniensia* 52(2), 240–59.

BNC 2014 *User manual = The British National Corpus 2014: User manual and reference guide (version 1.1) 2018*. http://corpora.lancs.ac.uk/bnc2014/doc/BNC2014manual.pdf (accessed 27 October 2022)

BNC Consortium 2007. *British National Corpus: XML edition*. Oxford Text Archive (accessed 27 October 2022).

Boas, Hans C. 2013. Cognitive Construction Grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Grammar*, 233–52. Oxford: Oxford University Press.

Brezina, Vaclav & William Platt. 2022. *#LancsBox X* 1.1.0 [software package].

Brezina, Vaclav, Abi Hawtin & Tony McEnery. 2021. The Written British National Corpus 2014 – Design and comparability. *Text & Talk* 41/5–6, 595–615.

Burnard, Lou. 2007. *BNC user reference guide*. www.natcorp.ox.ac.uk/docs/URG/BNCdes.html

Bybee, Joan L. 2013. Usage-based theory and exemplar representations of constructions. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Grammar*, 49–69. Oxford: Oxford University Press.

Cappelle, Bert. 2006. Particle placement and the case for 'allostructions'. In Doris Schönefeld (ed.), *Constructions*, Special volume 1: *Constructions all over: Case studies and theoretical implications*. urn:nbn:de:0009-4-6839. Available at www.constructions-online.de

Croft, William. 2001. *Radical Construction Grammar. Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Culicover, Peter W. 2013. *Grammar and complexity: Language at the intersection of competence and performance*. Oxford: Oxford University Press.

Davies, Mark. 2008–. *The Corpus of Contemporary American English* (COCA). www.english-corpora.org/coca/ (accessed: 20 April 2023).

Dehé, Nicole & Yordanka Kavalova. 2006. The syntax, pragmatics, and prosody of parenthetical *what*. *English Language and Linguistics* 10(2), 289–320.

De Vaere, Hilde, Julia Kolkmann & Thomas Belligh. 2020. Allostructions revisited. *Journal of Pragmatics* 170, 96–111.

Donaldson, James. 2013. On elliptical *why*-questions. MSc, The University of Edinburgh.

Fernández-Pena, Yolanda. 2021. Towards an empirical characterisation and a corpus-driven taxonomy of fragments in written contemporary English. *RÆL, Revistæelectrónica de Lingüística Aplicada* 20(1), 136–54.

Fernández-Pena, Yolanda & Javier Pérez-Guerra. 2024. Why *why*-fragments? A corpus-based constructionist analysis of their form and meaning. In Stefania Maci & Giovanni Garofalo (eds.), *Investigating discourse and texts: Corpus-assisted analytical perspectives*, 131–60. Lausanne: Peter Lang.

Fernández, Raquel, Jonathan Ginzburg & Shalom Lappin. 2007. Classifying non-sentential utterances in dialogue: A machine learning approach. *Computational Linguistics* 33(3), 397–427.

Fillmore, Charles J., Paul Kay & Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of 'let alone'. *Language* 64(3), 501–38.

Fox, John & Sanford Weisberg. 2019. *An R Companion to Applied Regression*. Thousand Oaks, CA: Sage.

Goldberg, Adele E. 2000. Patient arguments of causative verbs can be omitted: The role of Information Structure in argument distribution. *Language Sciences* 34, 503–24.

Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Goldberg, Adele E. 2013. Constructionist approaches. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Grammar*, 15–31. Oxford: Oxford University Press.

Goldberg, Adele E. 2019. *Explain me this. Creativity, competition, and the partial productivity of constructions*. Oxford: Oxford University Press.

Gordon, David & George Lakoff. 1971. Conversational postulates. In *Papers from the Seventh Regional Meeting, Chicago Linguistic Society, April 16–18, 1971*, 63–84. Chicago: University of Chicago.

Halliday, M. A. K. & Christiane Matthiessen. 2014. *Halliday's Introduction to Functional Grammar*, 4th edn. Abingdon: Routledge.

Harrell Jr, Frank E. 2022. *_rms: Regression Modeling Strategies_. R package version* 6.3-0. https://CRAN.R-project.org/package=rms

Heine, Bernd, Gunther Kaltenböck & Tania Kuteva. 2020. On the status of *wh*-exclamatives in English. *Functions of Language* 27(2), 207–33.

Huddleston, Rodney & Geoffrey K. Pullum *et al*. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Hugou, Vincent. 2017. The WHX construction (*what the hell…?*) and intensity: A corpus-based study. *Lexis: Journal in English Lexicology* 10, 1–30.

Johnson, David E. 1975. Why delete Tense? *Linguistic Inquiry* 6(3), 481–9.

Kim, Jong-Bok. 2017. *Why*-stripping in English and Korean: A Direct Interpretation approach. *Studies in Generative Grammar* 27(4), 733–51.

Kim, Jong-Bok & Anne Abeillé. 2019. *Why*-stripping in English: A corpus-based perspective. *Linguistic Research* 36(3), 365–87.

Kim, Jong-Bok & Mark Davies. 2019. English *what with* absolute constructions: A Construction-Grammar perspective. *English Language and Linguistics* 24(4), 637–66.

Kortmann, Bernd. 1991. *Free adjuncts and absolutes in English: Problems of control and interpretation*. Abingdon: Routledge.

Lambrecht, Knud. 1990. 'What, me worry?' – 'Mad Magazine Sentences' revisited. In Kira Hall (ed.), *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society. February 16–19, 1990*, 215–28. Berkeley, CA: Berkeley Linguistics Society.

Love, Robbie, Claire Dembry, Andrew Hardie, Vaclav Brezina & Tony McEnery. 2017. The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations. *International Journal of Corpus Linguistics* 22(3), 319–44.

Merchant, Jason. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27, 661–738.

Michaelis, Laura A. & Hanbing Feng. 2015. What is this, sarcastic syntax? In Kerstin Fischer & Kiki Nikiforidou (eds.), *On the interaction of constructions with register and genre*, 148–80. Amsterdam: John Benjamins.

Nakao, Chizuru, Masaya Yoshida & Ivan Ortega-Santos. 2012. On the syntax of *why*-stripping. In Nathan Arnett & Ryan Bennett (eds.), *Proceedings of the 30th West Coast Conference on Formal Linguistics*, 270–80. Somerville, MA: Cascadilla Proceedings Project.

Osawa, Satoko. 2011. The EPP and Mad Magazine sentences. *IVY* 44, 19–36.

Põldvere, Nele & Carita Paradis. 2019. Motivations and mechanisms for the development of the reactive *what*-x construction in spoken dialogue. *Journal of Pragmatics* 143, 65–84.

Põldvere, Nele & Carita Paradis. 2020. 'What and then a little robot brings it to you?' The reactive *what*-x construction in spoken dialogue. *English Language and Linguistics* 24(2), 307–32.

Progovac, Ljiljana. 2006. The syntax of nonsententials. Small clauses and phrases at the root. In Ljiljana Progovac, Kate Paesani, Eugenia Casielles-Suárez & Ellen Barton (eds.), *The syntax of nonsententials: Multidisciplinary perspectives*, 33–71. Amsterdam: John Benjamins.

Pskit, Wiktor. 2017. On 'small' structures in syntax: Small clauses, nonsententials, and small nominals. *Heteroglossia* 7, 123–33.

R Core Team. 2022. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. www.R-project.org

Siemund, Peter. 2015. Exclamative clauses in English and their relevance for theories of clause types. *Studies in Language* 39(3), 697–727.

Siemund, Peter. 2017. English exclamative clauses and interrogative degree modification. In Maria Napoli & Miriam Ravetto (eds.), *Exploring intensification: Synchronic, diachronic and cross-linguistic perspectives*, 209–30. Amsterdam: John Benjamins.

Siemund, Peter. 2018. *Speech acts and clause types. English in a cross-linguistic context*. Oxford: Oxford University Press.

Van de Velde, Freek, Hendrik De Smet & Lobke Ghesquière. 2015. On multiple source constructions in language change. In Hendrik De Smet, Lobke Ghesquière & Freek Van de Velde (eds.), *On multiple source constructions in language change*, 1–17. Amsterdam: John Benjamins.

Venables, W. N. & B. D. Ripley. 2002. *Modern Applied Statistics with S*. New York: Springer.

Weir, Andrew. 2014. *Why*-stripping targets Voice Phrase. In Hsin-lun Huang, Ethan Poole & Amanda Rysling (eds.), *NELS 43. Proceedings of the 43rd Meeting of the North East Linguistic Society*, 235–48. New York: The City University of New York.

Weir, Andrew. 2017. But write what? In Nicholas LaCara, Keir Moulton & Anne-Michelle Tessier (eds.), *A schrift to fest Kyle Johnson*, 401–8. Amherst, MA: University of Massachusetts.

Woods, Rebecca & Luis Vicente. 2021. Metacommunicative-*why* fragments as probes into the grammar of the speech act layer. *Glossa*: *A Journal of General Linguistics* 6(1), 84.

Yoshida, Masaya, Chizuru Nakao & Iván Ortega-Santos. 2015. The syntax of *why*-stripping. *Natural Language and Linguistic Theory* 33, 323–70.

Zaitsu, Anissa. 2018. Why make sense of silence? The clausal syntax of a reduced *why*-question. PhD dissertation, UC Santa Cruz.

Zaitsu, Anissa. 2020. Modality force and syntax in an understudied class of reduced *why*-questions in English. *Glossa: A Journal of General Linguistics* 5(1), 1–37.