



RESEARCH ARTICLE

Lexical *be*¹

Philip Miller¹  and Peter W. Culicover² 

¹Université Paris Cité, CLILLAC-ARP, F-75013 Paris, France

²The Ohio State University and University of Washington

Corresponding author: Philip Miller; Email: philip.miller@u-paris.fr

Received: 25 October 2024; **Revised:** 26 October 2024; **Accepted:** 31 October 2024

Keywords: Corpus linguistics; Directives; English auxiliaries; Grammatical constructions; Lexical BE

Abstract

We explore the surprising lexical *be* construction in English (e.g. *Why don't you be quiet?*). After an overview of previous discussions, an investigation of the use of lexical *be* in the COCA and SOAP corpora is provided. It is shown that its distribution is highly skewed and that it is completely felicitous only under a very limited set of conditions. An account of lexical *be* is then provided showing that the conditions that license it are inherited from more general constructions, most importantly the negative imperative construction and the '*Why don't you*' construction. In this light, it is suggested that the lexical *be* construction, with its special properties, provides strong evidence for a constructional approach to linguistic competence along the lines of Goldberg (1995), Culicover and Jackendoff (2005), Sag (2012).

1. Introduction

Two hallmarks of Geoff Pullum's remarkable career are (i) careful attention to the fine-grained properties of English constructions (as seen, for example, in Huddleston et al. (2002)) and (ii) compelling arguments for a model-theoretic approach to the licensing of grammatical form, as contrasted with familiar derivational approaches (e.g. Pullum and Scholz 2001). In this paper, we discuss an understudied set of constructions that nicely exemplify both of these threads, what Huddleston et al. (2002, 114) (H&P) call 'lexical *be*', as in (1).

- (1) a. Why don't you be quiet?
- b. Why don't you be the judge?

¹ This paper was presented at the 2019 congress of the SAES in Aix-en-Provence, at the 2020 LSA meeting in New Orleans and in Edinburgh on August 31 2023 at the conference 'English language and linguistic theory: A tribute to Geoff Pullum'. We would like to thank the audiences for useful questions and comments, as well as the anonymous reviewers for this journal.

Cases like (1) are striking because lexical *be* does not invert with the subject in the interrogative construction, but undergoes *do*-support. Thus, it contrasts sharply with copular *be* in interrogatives, which behaves like an auxiliary, as in (2).²

- (2) a. Why aren't you quiet?
b. Why aren't you the judge?

By comparison, the unacceptability of examples such as (3) shows that the distribution of lexical *be* is severely limited.

- (3) a. ??Why do you be the judge / quiet?
b. *Do you be the judge / quiet?
c. *You don't be the judge / quiet.

Our investigation into the properties of lexical *be* is organized as follows. In Section 2, we review prior discussions of lexical *be* in the traditional literature. These discussions do recognize the existence of lexical *be* but do not observe most of the details of the construction or its actual use.

These details are developed in Section 3 through a detailed corpus investigation. We show that the usage of lexical *be* in American English is highly skewed in its distribution and that it is completely felicitous only under a very limited set of conditions.

In Section 4, we propose to explain the distribution of lexical *be*. We show that the conditions that license lexical *be* in examples like (1) are inherited from more general constructions – most importantly, the negative imperative construction (NI; *Don't you be so pushy!*) and the '*Why don't you*' construction (WDY; *Why don't you sit down?*).

This brings us to the question of providing a grammatical account of the properties of lexical *be*. Collins (2006) proposes a derivational account of this construction, which we discuss in Section 5. Collins goes beyond the traditional literature and identifies some of the core properties of the construction. Nevertheless, we argue that while the derivational account stipulates that lexical *be* is possible, it does not account for its distribution or most of the fine-grained details.

Section 6 provides an account of lexical *be* in terms of a network of related constructions (cf., for example, Goldberg 1995; Sag 2012). Constructions state licensing conditions on correspondences between syntactic structure, phonological form and semantic interpretation; they exemplify the model-theoretical approach to grammar argued for by Pullum and Scholz (2001). A central property of constructional analyses is constructional inheritance: idiosyncratic constructions share certain licensing conditions of more general constructions.

Section 7 concludes with a summary of the more general theoretical implications of our account.

² In English linguistics, since Huddleston (1976), it is customary to use the term 'auxiliary' to mean 'verb that has the so-called NICE properties' (see Huddleston et al. (2002, 92ff), rather than in the traditional 'helping verb' sense. Under this definition, *be* is an auxiliary in English, even in its copular uses, since it can invert with the subject and take negative *not* without *do*-support ('Is Kim happy?'; 'Kim isn't happy'). The constructions we are investigating here are exceptional in that *be* behaves like a lexical verb (i.e. a verb that does *not* have the NICE properties and consequently exhibits *do*-support).

2. Prior discussion on lexical *be*

Very little has been said about lexical *be* in the traditional literature. The first mention³ we find is Erades (1953), who cites the following example.⁴

(4) Why don't you be reasonable?

Quoting Wood, Erades says that this is 'certainly not a solecism' and that 'its meaning is not the same as *why aren't you reasonable?*'. He further observes that *Why don't you* + inf. 'is quite frequently used as a kind of rhetorical question which is, in effect, not a question at all but a suggestion or a piece of advice'. For verbs other than *be*, he suggests, the structure is ambiguous, with only the context making the intent clear. Wood also comments on a change in the meaning of *be*: the verb itself has 'a slightly different meaning [...] suggestive of purpose, effort, or activity directed toward the specified end. *Be* means something like "take steps to become" or "bring yourself to be"'. He cites, for instance, a case like *Why don't you be a doctor?*. We see here the first recognition that lexical *be*, unlike auxiliary *be*, explicitly expresses a directive or desiderative force.⁵

Denison (2000, §3.3.1) points out some examples that are quite different from (1), such as (5) (his examples (50) and (52)):

- (5) a. Did littering the streets not once be considered breaking the law?
b. We agree that particular students do be flogged.

He also cites the classical cases, including (6) from Wharton's *Age of Innocence* (his example (57)), which he calls the 'the quasi-imperative *Why don't you be ...?* construction', which he claims is 'now fully accepted in standard'.

(6) Who's 'they'? Why don't you all get together and be 'they' yourselves?

He further states that 'The typical syntactic contexts for tensed DO + BE are non-assertive, frequently both negative and either interrogative, quasi-imperative, or conditional. If BE is lexical, it usually forms a nonstative group-verb with its complement'.⁶

³ Kjellmer 2009: 75, cites an earlier reference – namely, Dietrich 1949, who 'discusses a number of occurrences from England [...] of *do*-support with *be*, most of them of the type *Why don't you be ...?*'. We have not been able to access Dietrich's book.

⁴ It appears that *English Studies*, at the time, had a column edited by Erades where it was possible to ask questions about points of English syntax. Readers of the journal would send in their answers, and Erades would provide a summary in the following issue. The example in (4) appears to have been the topic of a query by a reader. Erades briefly cites a couple of contributions, which he rejects (e.g. that *don't* is a misprint for *won't*) and then provides in full a contribution from 'Dr. Wood' (presumably Frederick T. Wood), which we summarize here.

⁵ For completeness, we mention here various brief references to lexical *be* which add nothing of substance to Wood's discussion. Hirtle 1965: 31, cites the last example mentioned (without reference to Erades or Wood) and contrasts it with *Why aren't you a doctor?* He notes that with *do* '*be* is thrown beyond the moment of speaking and refers to the future' and that the sentence 'evokes the notion of becoming a doctor'. Joly and O'Kelly 1990: 189–190, apparently follows Hirtle (without citing him), replacing *doctor* by *engineer* in their pair of contrasting examples. The phenomenon is also mentioned by Green 1975, Becker 2004: 404, Payne 2013: 32, Siemund 2018: 303–304.

⁶ Examples like (5) appear to belong to a different register from usual standard English, possibly more formal and/or archaic. We do not investigate such cases further in this paper, as they do not share the characteristics of

Huddleston et al. (2002, 114) propose a detailed description of lexical *be*. They provide cases similar to (1) and note that they are ‘virtually restricted to the negative’ as shown by the unacceptability of the variants without *not*, as in (7).

(7) ??Why do you be the judge / quiet?

They also note that in cases like 1, the utterance pragmatically conveys a directive (\approx ‘Be the judge. / Be quiet.’). They even provide 3SG variants (‘Why doesn’t he be ...’). And they point out a second subtype of the construction with *if*, illustrated in (8), noting that it again conveys a directive (‘you/he should be quick’):

- (8) a. If you don’t be quick, you’ll lose.
b. If he doesn’t be quick, he’ll lose.

It is important to stress that the acceptability of such cases of lexical *be* is truly surprising. In typical cases, attempting to use supportive *do* with *be* is highly unacceptable to the point of being unintelligible, even when the predicate is easily interpreted as actional, as seen in (9).

- (9) a. ??When did you last be the judge?
b. ??Where do you usually be polite?
c. ??You didn’t be the judge last time.
d. ??He didn’t be quick, so he arrived late.

With stative predicates, attempting to use supportive *do* leads to even stronger unacceptability (even with stage level predicates, as in (10c) and (10d)).

- (10) a. *Does your car be red?
b. *She doesn’t be French.
c. *Did she be present at the meeting?
d. *This brand doesn’t currently be available.

The only substantive theoretically oriented discussion of lexical *be*, to our knowledge, is that of Collins (2006), couched in a Principles and Parameters perspective. He bases his discussion on a set of naturally occurring data collected on the internet through Google searches. He cites classical cases like (1) and (8) above (e.g. his (5), (12), (13), (14)), and he also mentions examples like the following, where lexical *be* occurs with past tense supportive *do* in an *if* conditional, and in a main clause polar interrogative, respectively:

- (11) If he didn’t be careful, he would almost feel like he was falling in love with her. [his (21)]
(12) ‘And did you be nice?’ Santa asked sharply. [his (23)]

standard lexical *be* that we describe below and are not found in our corpus data. Similarly, we do not deal with other nonstandard varieties that have been mentioned in the literature by, for example, Payne (2013, 31-2), who provides examples with inflected *be* (*bes*, *beed*, his (61) and (64)). See Montgomery and Mishoe (1999) and Hollenbaugh (2023) for further discussion of inflected *be* in nonstandard American English.

Collins' central claim is that the construction is only possible with agentive *be*. He notes that lexical *be* is never inflected in his variety of English (*'If he be's/bes/bees careful, ...'* [his (35b): 207]).⁷ He also notes that the construction 'seems more acceptable in the context of an implicit command', as in (13).

(13) I suggest that you don't be late next time. [his (47a)]

We take up Collins' account of lexical *be* in Section 5 in our discussion of the conclusions to be drawn from the possibility of lexical *be*.

Finally, Kjellmer (2009) provides a corpus-based investigation of lexical *be* in British English, based on the British National Corpus (BNC), the CobuildDirect Corpus and the internet. He points out the predominance of *Why don't you be XP?* and suggests (p. 81) an analogy with negative imperatives noting the semantic proximity between the two. He also notes (p. 83) the predominance of the present tense. But, beyond this, his examples of lexical *be* are mostly taken from the internet and consequently, as with Collins' data, do not offer a clear picture of the relative frequency of the various subconstructions.

Before closing this section, let us explicitly make clear that the negative imperative with *be* (e.g. *Don't you be too hard on yourself.*) is not an instance of lexical *be*. As discussed below at the end of section 4, *do* in the negative imperative does not mark tense and is hence not a genuine case of *do*-support, contrary to what is the case with lexical *be*.

3. Corpus investigation of lexical verb *be*

In order to get a more detailed and representative picture of lexical *be*, we investigated its use in the COCA and in the SOAP corpora.⁸ We used various queries such as [*do*] *n'* *t* [*pp**] *be* (any form of *do* followed by *n'* followed by any personal pronoun followed by *be*) which were designed so as to retrieve the great majority of occurrences.⁹

Complete list of queries:

[*do*] *n'* *t* [*pp**] *be*
 [*do*] *n'* *t* [*pp**] [*r**] *be*
 [*do*] [*pp**] *be*
 [*do*] [*pp**] [*r**] *be*

⁷ Collins finds occurrences of agentive *be* with *when* but apparently none with actual *do*-support; for example, *When you be careful enough to check during the install procedure ...* [his (16)].

⁸ Corpus of Contemporary American English (<https://www.english-corpora.org/coca/>), 600 million words (at the time when the data was collected, in late 2018) and Corpus of American Soap Operas (<https://www.english-corpora.org/soap/>), 100 million words. We initially investigated the COCA and found that, though the construction was used in all five registers, all of the examples from written registers were in fact cases of reported speech. The construction is thus typical of spoken English, which led us to choose the SOAP corpus, which provides typical conversational English, to obtain further data.

⁹ We did not check for non-pronominal third-person subjects. However, given that we found only 4 cases of pronominal third-person subjects (out of 208 cases) and that pronominal subjects are much more frequent than full NP subjects in ordinary conversational English (cf., for instance, Francis et al 1999, who found that 91% of the subjects in a part of the Switchboard Corpus, a corpus of telephone conversations, were pronominal), it is very likely that there are very few such cases to be found, if any at all. In the queries, square brackets around *do* recover all the inflected forms, [*pp**] matches all personal pronouns, and [*r**] represents any adverb, the only ones recurring with any frequency being *just* and *ever*.

Table 1. Frequency of subconstructions involving lexical *be* in COCA and SOAP

Clause type	Total	COCA	SOAP	Example
WDYB	170	34	136	Why don't you be the judge?
HDIB	19	17	2	How do I be the real me [...] ?
IYDB	12	8	4	if you don't be quiet you'll have to sit outside!
Polar-Int-Main	3	3	0	Do we be general or specific?
So-Adjunct	1	1	0	[...] so you don't be that controlling with him.
That-Subj	1	1	0	[...] demands that we don't be too materialistic
Main	2	1	1	You don't be.
Total	208	65	143	

[do] [pp*] not be
 [do] [pp*] not[r*] be
 [pp*] [do] n' t be
 [pp*] [do] n' t [r*] be
 [pp*] [do] be
 [pp*] [do] [r*] be
 [pp*] [do] not be
 [pp*] [do] not[r*] be

We eliminated noise and a certain number of cases which provided clear evidence of not being standard English. This provided a total of 208 (65/143)¹⁰ occurrences. Table 1 provides an overview of the results.

The most striking property of these data is that the distribution over construction types is extremely skewed. First, we see that the main clause *why*-interrogative subconstruction (WDYB) is by far the most frequent, comprising more than 80% of occurrences. Two further subconstructions occur comprising between 5 and 10% of occurrences, specifically (i) the *if*-conditional subconstruction (IYDB) identified by H&P (6% of occurrences)¹¹ and (ii) the main clause *how*-interrogative subconstruction (HDIB), not noticed by H&P, with 9% of occurrences, illustrated in (14).

- (14) a. How do I be a better person?
 b. How do you be friendly with a man who says he doesn't regret setting bombs?

The other cases found have a frequency close to 1%. Furthermore, numerous imaginable configurations were not attested at all (e.g. those in (9) above).

Second, we note that subjects are heavily constrained depending on the subconstruction, as shown in Table 2. Second-person *you* is present in 90% of the cases in the WDYB subconstruction and 75% in the IYDB subconstruction, whereas first and second person

¹⁰ After the total number of occurrences, we put in parentheses the number of occurrences from the COCA followed by the number from SOAP.

¹¹ For the *if*-conditional subconstruction, we have only counted occurrences with negation and supportive *do*.

Table 2. Frequency of subjects by subconstruction in COCA and SOAP

	WDYB	HDIB	IYDB	Other	Total
you	152	9	9	3	173
I	6	9	0	1	16
we	11	1	1	2	15
he	1	0	0	1	2
she	0	0	1	0	1
they	0	0	1	0	1
Total	170	19	12	8	208

appear to be balanced in the HDIB subconstruction. Third-person subjects are only marginally represented overall.

Third, polarity is constrained by subconstruction: all occurrences of the WDYB construction have *n't*, whereas all of the HDIB constructions are positive.¹²

Fourth, the form of *do* is highly constrained. It is always *do* or *don't* except for one occurrence of *doesn't* with the subject *she* and two of *did* with subject *he* and one of *didn't* with the subject *they*. Of special interest is the absence in our corpora of past tense uses with first- and second-person subjects, as in (15), despite the fact that the form makes perfect sense (viz., '{I/You} should have been quiet/the judge').

(15) Why didn't {I/you} be quiet/the judge?

As for speech acts, as suggested by H&P, directives dominate.¹³ This is centrally the case because the WDYB construction always clearly has this interpretation with a *you* or *we* subject, as in (1a) above. Similarly, most cases of the IYDB construction have clear directive force, as in (8) above.

Other cases, illustrated in (16), do not have actual directive force. However, all of these involve a deontic or desiderative stance that is closely linked to directive force, in the sense that the subject should be or wants to be the agent of the action, as indicated in the glosses.

- (16) a. If she doesn't be careful, she just might end up ... (≈ She should be careful.)
 b. Why don't I be the one to reach out to him (≈ I propose I should be the one to reach out to him.)
 c. How do I be like you? (≈ I want to be like you. How should I do that?)

Finally, Table 3 provides information about the category of the predicate in lexical *be* constructions. Clearly, AP and NP predicative complements strongly dominate because they can easily denote actional predicates. The low number of PP predicative complements can be

¹²Note that one occurrence of HDIB has negation in the infinitival VP: 'How do you not be in a position where you can legitimately be said to be anti-civil rights and are sabotaging the bill?'

¹³We found more than twice as many occurrences in the SOAP than in the COCA, despite the fact that the spoken section of the COCA (120 million words at the time when the data was collected) is larger than the SOAP. This is likely explained by the fact that directives are typical of interactional dialogue, which is the register of the SOAP, whereas the spoken section of the COCA includes, for instance, news bulletins and other forms of monologue.

Table 3. Category of the predicate of lexical *be*

Category	COCA	SOAP	Total
Adv	1	4	5
AP	23	68	91
No Compl	2	0	2
NP	30	67	98
PP	8	3	11
VP	1	1	2
Total	65	143	208

explained by the fact that typical PP locative complements are not actional (*??Why don't you be in the garden*). This can be compared with the actional PP predicate in our corpus example *Why don't you be more like me?* (\approx 'Why don't you behave more like me?').

4. Explaining the range of meanings and the differences in acceptability of lexical *be*

In this section, we tease out the fine-grained conditions that license the use of lexical *be*. The first point concerns the form of the clause (i.e. whether it is a *why/how*-interrogative or some other interrogative, an *if*-clause, etc.). The second concerns polarity, the third concerns the person and number of the subject, while the fourth concerns tense and aspect. The subtle properties of the various constructions involving lexical *be* all contribute to the argument for a constructional account, and against a derivational account such as Collins (2006).

We start with the most frequent construction, *Why don't you be XP* (WDYB), which accounts for over 80% of occurrences of lexical *be* in our corpus data. Our hypothesis is that this construction is a blend inheriting properties of the more general *Why don't you VP* construction (WDY) and the negative imperative with subject (NIS) construction *Don't NP VP*, where NP is usually *you*. Independently of lexical *be*, the more general WDY construction is widely used as a means of expressing weak directives (i.e. suggestions or invitations). With the imperative and deontic *should*, it is a member of a set of constructions expressing a gradient of decreasing deontic force.

- (17) a. Sit down!
 b. You should sit down.
 c. Why don't you sit down.

We suggest that the WDYB construction arises from the pressure to make the full gradient available to actional predicates naturally expressed with *be* as the verb:

- (18) a. Be specific about what you want.
 b. You should be specific about what you want.
 c. Why don't you be specific about what you want?

We also propose that this is facilitated by the fact that the sequence *don't you be* independently exists in the negative imperative with subject construction (NIS, for example, *Don't*

Table 4. Person and number of subjects in WDY and WDYB in the SOAP corpus

WDY			WDYB		
Why don't you do	569	76.7%	Why don't you be	152	89.4%
Why don't we do	148	20.0%	Why don't we be	11	6.5%
Why don't I do	21	2.8%	Why don't I be	6	3.5%
Why doesn't he do	2	0.3%	Why doesn't he be	1	0.6%
Why doesn't she do	0	0%	Why doesn't she be	0	0%
Why don't they do	2	0.3%	Why don't they be	0	0%
Why doesn't it do	0	0%	Why doesn't it be	0	0%
Total	742	100%		170	100%

you be late!). Blending the two constructions is further facilitated by their semantic similarity.

It is important to note that the gradient deontic force affects the way the directive is presented by the speaker to the addressee rather than the nature of the directive itself. As pointed out by Huddleston et al. (2002, 929ff), imperatives can convey a wide variety of directives (from commands and requests to advice, warnings, instructions, invitations, permission and wishes). It is, in fact, the properties of the directed action which determine how the directive is interpreted. In particular, if it is beneficial to the speaker, a command or request is most likely; if it is beneficial to the addressee, advice, warnings, instructions, invitations, permission and wishes are more likely.

The WDY construction can similarly express a broad range of directive meanings,¹⁴ but it presents them as not having the force of a command. In a case like *Why don't you be quiet!*, where the resulting situation is clearly beneficial to the speaker and not the addressee, the WDY construction presents, under the guise of an invitation, what can in fact only be understood as a command or request.

As stated above, our proposal that the WDYB construction results from a blend of the WDY and NIS constructions receives initial support from the presence of the negative auxiliary *don't*, which we claim is inherited from both the NIS and the WDY construction. The latter requires negation; positive *why* interrogatives do not have the directive force of WDY. Compare (17c) with *Why do you sit down?* Further support for the idea of a constructional blend can be obtained by a close examination of the subjects in WDY and WDYB. If WDYB was simply an extension of WDY, we would expect the range and proportions of pronominal subjects in both constructions to be the same.

A closer look at the data shows that this is not in fact the case. Consider Table 4:

In order to establish a useful comparison between WDYB and WDY, we have to eliminate cases where *Why don't you ... ?* is actually an information question (and hence not a case of WDY). A close look at the data shows that the information question interpretation typically arises with stative verbs (*Why don't you like them?*).

In this light, we searched the SOAP for occurrences of main verb *do* in WDY (i.e. the pattern *why do | does n' t PRON do*). Though it is in principle possible for such sequences

¹⁴Note that WDY allows a slightly narrower range of directive stances than the imperative, excluding wishes, where the event is not under control of the subject. Compare: *Sleep well.* and *#Why don't you sleep well?*

to denote information questions, an examination of the first 100 occurrences in the SOAP shows that they are all, in fact, relevant instances of weak directives – that is, cases of WDY. Further evidence that information question interpretations are not present here stems from the fact that third-person subjects are almost absent.

Both the WDY and the WDYB constructions show strong skewing toward second person and, to a lesser extent, toward first. However, it is interesting to note that there is a striking difference between the two in the distribution among first- and second-person pronouns. The bias toward second person is stronger for WDYB than for WDY.¹⁵ Since the weak directive interpretation is the same for both constructions, it alone does not explain this difference. We suggest that the bias toward second person in WDYB supports the hypothesis that it is a blend with NIS, where the subject is almost always *you*.¹⁶

Consider now the restriction to present tense we found in the WDYB construction. If WDYB were a simple extension of WDY, we would expect the same distribution of tenses in both constructions. But the data from the SOAP show that this is once again not the case. For WDY, the SOAP provides 74 occurrences of *Why didn't PRON do?*. This is unsurprising as these cases implicate a counterfactual deontic reproach interpretation similar to 'You should have' – for example,

(19) Why didn't you do this test before? (SOAP)

These 74 occurrences in the past tense are then clearly instances of WDY.

We thus obtain a proportion of 74 past tense occurrences of WDY for 742 occurrences in the present tense (see Table 4). This should be contrasted with 3 occurrences in the present tense of WDYB for 208 in the past tense (see discussion just before example (15)). These are clearly different proportions.

We suggest that the constraint on the present tense is inherited from the NIS construction where the form is always *don't* in the present. The rare cases where we find the past tense *didn't* can then be attributed to further ad hoc extensions of the construction, where the formal link to the negative imperative is overridden.¹⁷

Once lexical *be* is established in the WDYB construction, it can extend to further constructions where it keeps most of the same formal and semantic properties. Only two such constructions show any degree of productivity in our corpus data. Consider first the *If you don't be* (IYDB) construction, noted by Huddleston et al. (2002, 114). As they point out, it is used in contexts where the utterance indirectly conveys a directive as in:

(20) Look, if you don't be careful, this whole facade of yours is going to come crumbling down, (SOAP) (\approx You should be careful or ...)

¹⁵ Pearson's Chi-squared test gives a p-value = 0.001052, confirming that there is an interaction between person choice and WDY vs. WDYB. We thank Guillaume Desagulier for carrying out the statistical analysis.

¹⁶ The SOAP has only 47 occurrences of imperatives with indefinite pronoun subjects, of the type *Don't everyone talk at once*.

¹⁷ We leave open here and elsewhere the questions of whether and at what point ad hoc extensions of a construction are reinterpreted by speakers as a novel construction or as changes in the licensing conditions of the original construction.

The IYDB construction thus preserves the central semantic property of WDYB. It also preserves the formal properties: second-person subjects represent 75% of occurrences, and all occurrences are in the present tense and negative with the form *don't*.

Second, consider the *How do I/you be* (HDIB) construction (see also above examples (14)):

- (21) a. How do I be the real me [...]? (COCA)
 b. How do you be a mother to four children with everything else you do? (COCA)

From a semantic point of view, these examples appear to have diverged a little with respect to the directive meaning of the WDYB construction. There is no actual directive expressed. Rather, the occurrences express deontic modality ((21a) \approx 'I should be the real me') often combined with an element of desiderative modality ('I want to be the real me'). Formally, the construction shares *wh*-interrogative status with the WDYB construction. Furthermore, all of the examples are in the present tense with the form *do*. However, there is a significant change with respect to subjects. First, 10 of the 19 examples are first person. But beyond that, the 9 occurrences of second person do not have the same interpretation as in WDYB and IYDB: rather than referring to the addressee, they clearly receive a generic interpretation, applicable to the case of the speaker.¹⁸ Clearly, the speaker in (21b) is generalizing over her own desire to be a (good) mother. There is no attempt to act on an addressee (this explains our choice of I in the HDIB acronym, rather than Y as in WDYB and IYDB). This change in interpretation with respect to the subject is easily understood given the shift from directive (influence on the addressee) to deontic/desiderative (attempt to influence oneself).¹⁹

Four additional cases appear in the corpus, comprising 7 occurrences. These diverge from the central constructions that we have been discussing. They appear in Table 1 and include the examples in (22).

- (22) a. should I try to help save these lives or do I help my three children or do I be with my father on his death bed?
 b. The relative smallness of the Westwood town houses, says resident Vance Reese, demands that we don't be too materialistic.
 c. So, you have to work with him around that so you don't be that controlling with him.
 d. I'm interested in what you think being dead would be like. [...] You don't think. You don't remember. You don't be.

Interestingly, the first three of these express a deontic meaning. Note specifically the deontic *should* in the initial clause of (22a). Clearly, the two following clauses should be

¹⁸ This finding led us to check the COCA and SOAP for occurrences of generic *one* subjects in the HDIB construction, *how does one be ... ?*. We could not find any occurrences, perhaps because of the register clash between the formality of generic *one* and the typically spoken register of the lexical *be* constructions. Cases of this type are easily found on the web, however – for example, *How does one be oneself in the pulpit?* (<https://scholar.csl.edu/cchom2/47/>).

¹⁹ We should note the presence of a single example of HDIB in the SOAP which appears to require interpretation as a simple question of fact, without any deontic or desiderative (much less directive) stance – namely, *how do you be so heartless?*. Further, contrary to typical cases of HDIB, here, *you* denotes the addressee. To our ear, this example sounds much less natural than the examples of WDYB, IYDB and HDIB that we have been discussing.

understood as expressing the same modal judgment. Note also that all of these have first- and second-person subjects and are in the present tense with the form *do*, following the preferences found in the three central constructions. Example (22d), where there is clearly no deontic component present, seems to illustrate a different situation which can enhance the acceptability of lexical *be* – namely, cases where there is no other simple and obvious way of conveying precisely the same content. As we have seen, lexical *be* is restricted to actional predicates. Here, without a predicate, it seems to keep this actional meaning, making it preferable to the obvious alternatives (*You aren't / You don't exist*) which would tend to receive a strictly stative interpretation.

Similarly, it is interesting to note that the single instance of the WDYB construction with past tense *didn't* found in our corpus is clearly a case of sequence of tense in free indirect discourse:

- (23) Laughter floated down the passage, the stilted sound of people ill at ease. Tamborel made a face. Why didn't they just be quiet?

Clearly, the internal speech of the protagonist would be the standard *Why don't you just be quiet?*

It is, of course, possible to find occurrences of WDYB with a past tense that is not the result of sequence of tense by randomly searching the web, even in the third person (though these are clearly much less frequent than second person) for instance:

- (24) a. Why didn't you just be a teacher like your mother or your grandmother?²⁰
 b. Once Cat let the Kingslayer go, why didn't he just be cooperative and let Brienne guide him to KL?²¹

More generally, other types of *wh*-questions can be found (though they are clearly rare, and some occurrences are obviously produced by non-native speakers). These seem much more acceptable when they can be interpreted in context as involving a deontic stance.²²

- (25) a. [...] the biggest challenge is to sort of sit down and ask yourself what your role is in this space. So when do you be an active citizen? And when do you just sit down and keep quiet?²³
 b. I'm dating Chris Winters right now and I'm on level 13 and I went on all of the dates that pop up of him so when does he actually become mine like it says in the profile dating coming soon ... when do I actually be his girlfriend?²⁴

²⁰ <https://www.phillipwarfield.com/blogs/why>

²¹ <https://asoiaf.westeros.org/index.php?/topic/88302-why-did-jaime-resist-brienne/>

²² This explains why lexical *be* is much less acceptable in reason questions with *how come* than in *why* questions: the former are never used to convey a directive. Interestingly, the COCA does contain an example of apparent lexical *be* in a *how come* question, '*How come you don't be the new preacher?*' but the context shows that this is Jamaican English, and we have excluded it from our data as nonstandard. Obviously, it would be easy to account for such variation within the constructional framework we adopt here, but this goes beyond the scope of the present paper.

²³ <https://www.nature.com/articles/d41586-022-01149-5>

²⁴ <https://www.chaptercheats.com/qna/android/273566/Hollywood-U-Rising-Stars-Answers.htm?qid=110230>

In (25a), which comes from a paper published in a scientific journal, there is a clear deontic stance that is inferrable from the context (note the importance of *your role* and the ease of paraphrasing with *should*: *So when should you be an active citizen?*), and the example feels very acceptable. In (25b), however, there is no deontic stance at all. The utterance is a strict information question (that receives answers in the discussion forum about the dating game being referred to). The example seems less acceptable to us, and there is independent evidence of sloppy writing in the post (*the dates that pop up of him*). It is possible that *be* is used here instead of *become*.

In sum, if one conducts random searches on the web in general, one can find examples that diverge to varying degrees in meaning and form from the three central WDYB, IYDB and HDIB constructions we have discussed. The advantage of working on a well-defined corpus, as we have done here, is that it provides some idea of which cases are in fact relatively frequent collocations and which are much less frequent. Our proposal is that once lexical *be* is established in the three central constructions, with their related semantic and formal properties reviewed above, it can give rise to ad hoc extensions. The closer such extensions are to the three central constructions in their semantic and formal properties, the more likely they will be to be produced and the more natural they will sound.

Throughout this discussion, we have assumed that lexical *be* is only possible for copular uses of *be*. The following example from the COCA suggests that passive helping verb *be* can be lexical:

- (26) But the question becomes, when you start to see signs of this, and there were pervasive signs, clear signs, just a couple months ago, why did he not be declared unfit? Why did he have that clearance? (COCA)

This is clearly a true passive and not a copular construction. *Be* here is not agentive, it is meaningless (the corresponding active without *be* is truth-conditionally equivalent: *Why did someone not declare him unfit?*). The sentence is acceptable because it has clear deontic force (\approx 'he should have been declared unfit'), despite the fact that the agent who is the target of that force is left unexpressed in the passive.²⁵

This leads us to a final issue concerning the occurrence of *be* in imperatives. Since *be* systematically requires supportive *do* in the negative imperative construction (NI), should we consider that *be* is always lexical in imperatives? Although they are marginal, there are cases of the passive (and much more rarely the progressive) auxiliaries in imperatives, as in (27).

- (27) a. Be warned/advised/seated/assured ... (all well attested in the COCA)
b. I'm sending you the audio file. Be looking for it. (COCA)

When negated these also require *do*, and some of these clearly allow readings as actual passives rather than as copular clauses.

²⁵ Note that Collins (2006) stars a structurally similar sentence: **If you don't be seen, you will escape* (203, his (2b)). Our position is that the sentence is grammatical but unacceptable, because it doesn't exhibit the typical deontic reading associated with IYDB. Under the usual IYDB interpretation, it should mean 'You should be seen or you will escape', which does not make sense.

- (28) a. Don't be fooled/intimidated/discouraged/alarmed/misled/swayed/caught.
 b. Don't be telling me what to do. (COCA)

And *be* occurs as well with emphatic *do*.

- (29) Do be seated.

The difference between the imperative and WDY is that the occurrence of *do* in the imperative does not mark tense, and hence is not a genuine case of *do*-support; it is in the tensed question WDY. *Do* in the imperative is tenseless, as shown by the fact that there is no agreement with a third-person singular subject, as in (30) (see Culicover 1971).

- (30) a. Don't (*Doesn't) anyone move! (COCA)
 b. Don't (*Doesn't) everyone jump to conclusions! (COCA)

So bare *be* occurs in the two constructions but for different reasons, and it satisfies different interpretative constraints in the two constructions. In the imperative, *be* must be understood as 'active' (Partee 1977), and as directly or indirectly under the control of the addressee. In WDY, it has the feature of being under the direct control of the subject.

We conclude that lexical *be* is licensed by the three clearly productive constructions WDYB, IYDB and HDIB. Other uses are ad hoc extensions of these that are more likely to be produced and to be acceptable if they conform to the semantic and formal properties we have uncovered, especially the presence of deontic force.

5. Collins' minimalist approach to lexical *be*

We consider now the question of how to account for lexical *be* in a grammar. The only account in the contemporary generative literature we are aware of is that of Collins (2006). Collins proposes to account for lexical *be* by assuming that there are two verbs *be*: auxiliary *be* and agentive *be*. Following Chomsky (1995), he argues that only semantically vacuous verbs can move to Infl in English. On the assumption that auxiliaries are semantically vacuous, this provides an account for the fact that they do not show *do*-support in the NICE constructions. In a nutshell, Collins proposes that agentive *be* has an actional meaning that makes it unable to raise to Infl. Thus, it requires *do*-support because it is not semantically vacuous.

We argue that this approach is unsatisfactory for three reasons. First, we show in Section 5.1 that the agentive/ non-agentive contrast is not a property of *be*, but of the predicative complement (see Rothstein 1999). Second, we argue in Section 5.2 that the more general derivational approach lacks explanatory value. Third, we show in Section 5.3 that the derivational approach is not able to account for the systematic differences in frequency and acceptability between the different uses of lexical *be* observed in Section 4.

5.1. Should one distinguish agentive and non-agentive *be*?

Is the distinction between agentive and non-agentive *be* warranted? Rothstein (1999) argues against the assumption that there is more than one copular verb *be* (contra Partee 1977 and

Dowty 1979). We follow Rothstein here in claiming that differences in agentivity with *be* are due entirely to the properties of the predicative complement. A central piece of evidence for this is that the relevant differences in interpretation appear in various constructions that involve a subject–predicate relation without any overt verb at all. This can be seen by noting the parallelism between the a. and b. versions of the following pairs, independently of the presence or absence of the copula.

- (31) a. Sandy tried to be quiet / brave.
b. #Sandy tried to be tall / French.
- (32) a. Sandy always strives for brave over cautious.
b. #Sandy always strives for tall over short.
- (33) a. boys, I want to see you brave and manly, and I also want to see you gentle and tender. (COCA).
b. #boys, I want to see you tall and French.
- (34) a. all she wanted was to be in bed between the coolness of the sheets, with Max quiet in her arms (COCA)
b. # with Max tall / French in her arms

The subject control construction with *try* in (31) implicates that the subject is the intentional agent of the predicate – hence the felicity of adjectival predicates which allow an agentive reading like *quiet* and *brave* and the incompatibility with ILP predicates like *tall* and *French*. Similarly, the implied goal-directed action in (32) is embodied in the predicates without *be*. Despite the absence of *be*, the NP–AP complementation of the perception verb in (33) and the *with NP XP* construction in (34) allow predicates with an agentive reading but not ILPs.²⁶ It thus appears that the relevant interpretative differences are properties of the predicates and that the range of possible predicates is constrained by each construction. Ockham’s razor suggests that there is no reason to assume that different versions of *be* are in any way required to account for them.

In this light, let us have a closer look at the central data on which Partee (1977) bases her argument in favor of distinguishing a specific agentive *be* (her (50)–(57), 306).

- (35) a. John is noisy. (John makes a lot of noise.)
b. John is being noisy. (John is making a lot of noise.)
c. The river is noisy. (The river makes a lot of noise.)
d. *The river is being noisy. (The river is making a lot of noise.)

Partee argues that there is an agentive *be* that requires a volitional subject. This is clearly not a stative verb, as evidenced by its compatibility with the progressive in (35b). The volitionality condition explains the impossibility of (35d), where the subject is inanimate (Partee uses the variants with *makes/is making* to establish that volitionality is not simply a property of the progressive, but is linked to *be* itself).

²⁶ We are not claiming that the class of predicates that are felicitous in the NP–AP complementation of perception verbs and the *with NP XP* construction is the same as that of actional predicates which are acceptable with lexical *be*. The former constructions allow a much broader range of predicates – namely, (apparently) Stage Level Predicates.

We argue that even if one assumes that Partee is correct about the need to distinguish a specific agentive *be*, it would be irrelevant to our discussion because that *be* would be an auxiliary, rather than a lexical verb. It is impossible to use the NICE constructions to see this if one starts from (35b) because of the presence of the progressive auxiliary. But if we put the sentence in the past tense, both the property reading of (35a) and the agentive reading of (35b) are available:

- (36) a. As a child, John was noisy.
 b. During the trip home, John was intentionally noisy so his father told him to calm down.

Now consider what happens if we try to apply negation or subject auxiliary inversion to (36b):

- (37) a. During the trip home, John intentionally wasn't noisy.
 (compare: ??During the trip home, John intentionally didn't be noisy.)
 b. Was John intentionally noisy during the trip home?
 (Compare: ??Did John intentionally be noisy during the trip home?)

Clearly, the most natural step is to treat *be* as an auxiliary and simply add *n't* and invert as appropriate. Note that this is the case even under the volitional reading of these sentences, where *be* is clearly Partee's agentive *be*. The variants with *do*-support are similar to the cases of lexical *be* given in (22) cited earlier, but they are much less acceptable given the lack of any deontic stance.

We conclude that Partee's arguments do not bear on our discussion of lexical *be* because, in most cases, there is no reason to consider her agentive *be* as anything other than an auxiliary from a morphosyntactic point of view.

Beyond these considerations, Collins' proposal fails to account for the fact that in many cases, the auxiliary version of *be*, with the associated raising to Infl, has the same interpretation as the lexical version. For example, (38a) can have the same agentive interpretation as (38b).

- (38) a. If you aren't quiet, bad things will happen.
 b. If you don't be quiet, bad things will happen.

Examples like (38a) show that with appropriate predicates, such as *quiet*, either the non-agentive or the agentive interpretation is possible. Only when there is *do*-support is the agentive interpretation required. If agentivity resides in *be*, it should not be vacuous in (38a) under the agentive interpretation and should consequently not be able to raise over negation, contrary to fact.

5.2. Does the minimalist account actually explain anything?

Collins (2006, 212) explains the possibility of *do*-support with *be* by invoking a proposal of Chomsky (1995): in English, a verb with semantic content does not raise to Infl in NICE constructions. By assumption, auxiliary verbs do not have semantic content, so they are the only verbs that raise to Infl in English. Under standard assumptions, if Infl does not have a verb attached to it, *do* is inserted. Thus, we get alternations such as (39), where Infl is realized as Pres(ent Tense).

- (39) a. Sandy Pres speak French \Rightarrow Sandy speaks French.
 b. Sandy Pres *neg* speak French
 \Rightarrow Sandy doesn't speak French.
 \nRightarrow *Sandy speaks not French.
 c. Sandy Pres (*neg*) be tall \Rightarrow Sandy is(n't) tall.
 \nRightarrow *Sandy doesn't be tall.

Presumably, lexical *be* has semantic content on Collins's account because it selects an agentive subject. So it does not raise to Infl, triggering *do*-support in these constructions. The hypothesized difference between auxiliary and lexical *be* is illustrated in (40), paralleling (39).

- (40) a. If you Pres *neg* be_{aux} quiet \Rightarrow If you aren't quiet
 b. If you Pres *neg* be_{lex} quiet, bad things will happen.
 \Rightarrow If you don't be quiet, bad things will happen.

Besides the fact that this proposal is simply a notational device for marking the auxiliary/lexical distinction, it has a number of other shortcomings. First, Collins has to make a number of otherwise unmotivated ad hoc assumptions about empty heads and verb raising to get the correspondences between morphosyntactic form and meaning to come out right. Since this is a characteristic of virtually all analyses that assume empty heads and verb raising in order to derive surface forms, we do not dwell on it here; for extensive discussion, see Culicover and Varaschin (*In press*).

Second, Collin's proposal is falsified by the fact that English *have*, which functions both as an auxiliary and a lexical verb, can show NICE properties even when it expresses possession (and consequently is not semantically vacuous). The examples in (41) show that possessive *have* can participate in subject Aux inversion and precede negation.

- (41) a. Have you any $\left\{ \begin{array}{l} \text{idea what you're doing} \\ \text{suggestions} \end{array} \right\}$?
 b. I haven't $\left\{ \begin{array}{l} \text{a clue} \\ \text{any money right now} \end{array} \right\}$.

But the variants with *do*-support are equally well-formed.

- (42) a. Do you have any $\left\{ \begin{array}{l} \text{idea what you're doing} \\ \text{suggestions} \end{array} \right\}$?
 b. I don't have $\left\{ \begin{array}{l} \text{a clue} \\ \text{any money right now} \end{array} \right\}$.

It is not plausible to claim that *have* is semantically vacuous in (41) but not in (42), since the sentences are synonymous.

Similarly, the claim that modal auxiliaries are semantically vacuous also requires a neological understanding of semantic vacuity. Compare the two following sentences:

- (43) a. Kim may not be at home.
 b. Kim doesn't seem to be at home.

Both *may* and *seem* express a second order modal judgment (epistemic for *may*, evidential for *seem*). There is no reason to claim that *may* is semantically vacuous while *seem* is not. Both for *have* and for verbs expressing modality, such a claim would involve turning the morphosyntactic distinction between auxiliary and lexical into an unmotivated semantic distinction between vacuous and non-vacuous.

5.3. Summary

In summary, the derivational analysis of Collins (2006) does not provide a satisfactory account of the attested facts. The analysis predicts that all uses of lexical *be* with *do*-support will be grammatical when the interpretation is agentive. As we have seen, lexical *be* occurs in a range of constructions that are related in terms of meaning and idiosyncratic properties. Moreover, different uses of lexical *be* differ in frequency and acceptability. Furthermore, Collins puts aside (p. 211) the most frequent subconstruction, WDYB (*Why don't you be quiet?*). And his analysis cannot account for the fact that a deontic interpretation is required – he simply states that the construction ‘seems more acceptable in the context of an implicit command’ (p. 209). We conclude that Collins’ analysis is inadequate on both theoretical and empirical grounds.

6. A constructional analysis of lexical *be*

We turn now to our constructional analysis. Section 6.1 provides a brief overview of the formalism, and Section 6.2 spells out the particular constructions that license lexical *be* and their inheritance relationships.

6.1. Constructions

A grammar comprised of constructions exemplifies the model-theoretic approach to grammar of Pullum and Scholz (2001). We assume that the formalization of a construction has (at least) three corresponding TIERS (Jackendoff 2002; Culicover and Jackendoff 2005; Culicover 2021; Culicover and Varaschin In press): phonology (PHON), syntax (SYN) and meaning (SEM). An individual linguistic expression – a CONSTRUCT – is LICENSED by a set of constructions if it satisfies the LICENSING CONDITIONS imposed by the constructions. These conditions constrain the possible correspondences between the terms in the tiers.²⁷

On this view, a single lexical item is a construction. The construction for the verb *kick* is given in (44). For simplicity of representation, we use the spelling or the index of a constituent in PHON rather than the phonetic form. Cosubscripting indicates the correspondences.

$$(44) \left[\begin{array}{l} \text{PHON } \text{kick}_1 \\ \text{SYN } \text{V}_1 \\ \text{SEM } [\lambda y.\lambda x.\text{kick}'(\text{AGENT} : x, \text{PATIENT} : y)]_1 \end{array} \right]$$

²⁷ A construct does not have to satisfy all of the constructions in a grammar. Rather, all of the correspondences exemplified in a given construction must be licensed.

This construction says that *kick* is licensed in a linguistic expression if the phonetic form of the expression is /kɪk/ and it has the appropriate meaning. By the same token, a construct with the PHON representation /kɪk/ and the SEM representation $\lambda y.\lambda x.k\dot{i}s s'$ (AGENT:x,PATIENT:y)₁) is not licensed because *kick* does not mean ‘kiss’.

A somewhat more general construction that also has some idiosyncrasies licenses the idiomatic VP *sell NP down the river* is (45). \oplus signals ‘immediately precedes’.

(45) *sell NP down the river*

$$\left[\begin{array}{l} \text{PHON} \quad [\text{sell}_1 \oplus 2 \oplus \text{down}_5 \oplus \text{the}_6 \oplus \text{river}_7]_4 \\ \text{SYN} \quad [\text{VP } V_1, \text{NP}_2[\text{PP } P_5, [\text{NP } \text{Det}_6, \text{N}_7]_8]_3]_4 \\ \text{SEM} \quad \lambda x[\text{betray}'(\text{AG} : x, \text{TH} : 2')]_4 \end{array} \right]$$

Strictly speaking, the PHON tier does not have to be specified here, since the ordering is licensed by the HEAD-INITIAL VP construction in (46) below. And in fact, other VP orders are possible, for example, with heavy NP shift (*They sold down the river* [_{NP} *all the people who depended on them*]).

Other constructions state more general correspondences. For example, the English VP is licensed by (46). This construction says that V precedes the arguments and adjuncts of VP. The interpretation is left unspecified, since it is a matter of the particular lexical constructions such as (44) and (45). \ll signals ‘precedes’.

(46) HEAD-INITIAL VP

$$\left[\begin{array}{l} \text{PHON} \quad 1 \ll 2 \\ \text{SYN} \quad [\text{VP } V_1, \text{YP}_2] \end{array} \right]$$

6.2. Analysis

Now back to lexical *be*. We characterize the contexts in which lexical *be* appears in terms of the constructions that license it. There are two constructions to take into account, NEGATIVE IMPERATIVE and the *why don't you* construction WDY. NEGATIVE IMPERATIVE licenses tenseless *don't be* with directive force, and tensed WDY licenses weak directive force. The tensed WDYB construction is parasitic on these two constructions, allowing tensed *don't be* with weak directive force. Further extensions of the construction license the subconstructions IYDB and HDIB.

The details of the negative imperative construction are given in (47).

(47) NEGATIVE IMPERATIVE

$$\left[\begin{array}{l} \text{PHON} \quad [\text{don}'t_1 \oplus (2) \oplus 3]_5 \\ \text{SYN} \quad \left[\text{s}(\text{NP}_2), \text{AUX}[\text{NEG}]_1, [\text{VP } V[\text{BARE}]_4, \dots]_3 \right]_5 \\ \text{SEM} \quad [\text{DIRECTIVE}(\text{NEG}_1(4'(\text{AGENT} : \text{Addressee}|2', \dots)))]_5 \end{array} \right]$$

(47) states that in the negative imperative, there is *do*-support with the form *don't*, followed by an optional subject – *you* or a general addressee such as *anyone/everyone* – followed by the VP. The special case with the optional subject is what we have referred to as NIS. The

position of the subject aside, we do not need to stipulate the linear ordering properties of PHON, since they are licensed by independent constructions, such as (46).²⁸

Crucially, the negative imperative requires its own construction, since its morphosyntactic properties are idiosyncratic and do not follow from those that govern the imperative – it is a syntactic nut in the sense of Culicover (1999). The licensing of *don't* in this construction is an idiosyncrasy, since the imperative is tenseless and thus does not license productive *do*-support, as already noted (see (30)).²⁹

The WDY construction must stipulate the exact form *why don't you* as well as directive force.

- (48) WDY
- | | |
|------|--|
| PHON | $[\text{why}_5 \oplus \text{don't}_1 \oplus \text{you}_2 \oplus 3]_6$ |
| SYN | $[_S \text{NP}_2, \text{AUX}[\text{NEG}]_1, [_{VP} \text{V}[\text{BARE}_4, \dots]_3, \text{ADV}_5]_6]$ |
| SEM | $[\text{DIRECTIVE}(4'(\text{AGENT} : \text{Addressee}, \dots))]_6$ |

This is all simply bookkeeping, in a sense – keeping track of the unpredictable small details. But these formulations are necessary, since they take account of the irreducible idiosyncrasies of the language. They are, again, ‘syntactic nuts’, in the sense of Culicover (1999).

What is critical from the perspective of lexical *be* is that these two constructions share three properties: they have directive force, they license the form *don't*, and the verb has the bare form. The bare form is an idiosyncratic property of the imperative and negative imperative, and appears in the VP when there is an overt auxiliary *do*, as in the WDY construction. The robust correspondence between these morphosyntactic and semantic properties is recruited by other constructions that share these properties, thereby licensing *don't be* in a broader set of syntactic contexts.

Recall that on the present analysis, there is no construction that specifically licenses agentive *be*. We assume, however, that the licensing conditions for using the bare form with *don't* to express directive force in the well-established constructions (47) and (48) of English are available to speakers to express directive force in other contexts. The possibility of bare *don't be* in the negative imperative and the weak directive force of WDY together contribute to the novel construction WHY DON'T YOU BE that inherits the licensing conditions of (47) and (48). This construction is stated in (49).

- (49) WDYB
- | | |
|------|---|
| PHON | $[\text{why}_5 \oplus \text{don't} \oplus \text{you}_2 \oplus \text{be}_4 \dots]_6$ |
| SYN | $[_S \text{NP}_2, \text{AUX}[\text{NEG}]_1, [_{VP} \text{V}[\text{BARE}]_4, \dots]_3, \text{ADV}_5]_6]$ |
| SEM | $[\text{DIRECTIVE}(\text{be}'_4(\text{AGENT} : \text{Addressee}, \dots))]_6]$ |

Further extensions of (47) and (49) preserve the weak directive force but relax or modify one or more licensing conditions of the source constructions; the results are IYDB (50) and HDIB (51). We do not attempt to specify the interpretive details.

²⁸ Directive force in the imperative typically implies that the subject is agentive. However, this implicature does not apply to very weak directives – for example, *Sleep well!* or *Please be aware that S*.

²⁹ As discussed in Culicover (2021), the occurrence of *do*-support in imperatives is due to economy – in particular, maximization of constructional relatedness.

$$(50) \text{ IYDB} \left[\begin{array}{l} \text{PHON} \quad [\text{if}_5 \oplus \text{you}_2 \oplus \text{don't}_1 \oplus \text{be}_4 \dots]_6 \\ \text{SYN} \quad [\text{sC}_5, \text{NP}_2, \text{AUX}[\text{NEG}]_1, [\text{VP} \text{V}[\text{BARE}]_4, \dots]_3]_6 \\ \text{SEM} \quad \text{DIRECTIVE}(6') \end{array} \right]$$

$$(51) \text{ HDIB} \left[\begin{array}{l} \text{PHON} \quad [\text{how}_5 \oplus \text{do}_1 \oplus \{1/\text{you}\}_2 \oplus \text{be}_4 \dots]_6 \\ \text{SYN} \quad [\text{sNP}_2, \text{AUX}_1, [\text{VP} \text{V}[\text{BARE}]_4, \dots]_3, \text{ADV}_5]_6 \\ \text{SEM} \quad \text{DIRECTIVE}(6') \end{array} \right]$$

Ad hoc extensions of the directive force of these constructions allow for the occurrence of cases such as *why don't I/we/they be ...*; *why doesn't she/he be ...*, and even *why didn't I/you/she be ...*, *if she/he doesn't be ...*, *how does one be ...*, etc. The extensions that are judged more acceptable are those that preserve more properties of the prototypical cases, and clearly support a directive interpretation.

The approach sketched out here goes beyond a derivational analysis such as Collins (2006) because it explicitly addresses the correspondences between form and meaning at a level of detail where the actual phenomena can be accounted for. The capacity to refer to fine-grained licensing conditions is a hallmark of the constructional approach. In contrast, Collins' analysis predicts that all uses of lexical *be* that satisfy the two constraints mentioned above (agentive predicate and no inflection) will be equally acceptable.

Moreover, a constructional account is compatible with the view that judgments of acceptability reflect not only satisfaction of formal licensing conditions imposed by the grammar – that is, the set of constructions – but frequency, usage and constructional similarity (see, for example, Bybee 2010, 2013; Culicover et al. 2022). So, for example, the question *why didn't he be more careful* does not satisfy the explicit condition of (48) that the agent be the addressee and that the form be *don't*. Extension of the directive force with respect to a third person in the past is marginally possible, as illustrated by the corpus examples in (24) above. The relative difficulty of constructing a plausible context accounts for both the judgment of unacceptability out of context and the very rare occurrence of such a form.

7. Theoretical implications

To conclude, let us take a step back and consider briefly the theoretical implications of lexical *be*. One might, in fact, wonder why it is worth studying such a minor construction, since its frequency in spoken English is on the order of one occurrence per million words. We believe that lexical *be* is significant because it constitutes a classical instance of a 'syntactic nut'. That is, it is an idiosyncratic construction that licenses a highly restricted form with a certain interpretation or function.

While syntactic nuts are idiomatic, they may bear close relationships to the more general and productive constructions of the language,³⁰ as we have argued for WDYB, IYDB and

³⁰ Potts 2023 makes a similar argument in his discussion of the English Preposing in PP construction in this volume.

HDIB. A very different case that illustrates this point particularly clearly is the acceptable (52). What is idiosyncratic in this construction is that the direct object precedes rather than follows the verb, contrary to the general licensing conditions of the English VP in (46).³¹

(52) One swallow does not a summer make.

It is possible to adapt the pattern of (52) to novel situations, with a distinctive rhetorical force and gradient acceptability. For example,

- (53) a. ?One example does not a theory prove.
 b. ??One counterexample does not a theory disprove.
 c. ???One poorly constructed experiment does not an entire research program invalidate.

Given the acceptability of (52), examples like those in (53) can become more acceptable with repetition, or if they are understood as wordplay. Of course, corresponding examples in which the verb precedes the direct object are all undistinguished.

Constructions such as these are theoretically significant because speakers have clear intuitions about their acceptability, although they can be extremely idiosyncratic. The challenge is to explain how they fit into grammatical theory. It is not sufficient to claim that syntactic nuts like the lexical *be* construction are not part of the ‘core’ but rather belong to the ‘periphery’. Indeed, these constructions raise the typical learnability problems that are claimed to characterize core phenomena. Specifically, in the absence of negative evidence, there is no reason for a child exposed to (1a) not to simply assume that *be* can be used as a lexical verb with actional predicates in general, which would lead her to conclude that all the cases in (9) are perfectly well formed. Why is it then the case that speakers do not produce the full range of possibilities?

This approach rests on the view that judgments of acceptability reflect more than satisfaction of the licensing conditions imposed by the grammatical constructions. The constructions require that an expression satisfy not only narrow licensing conditions on syntactic form, but also conditions governing correspondences between syntactic form and phonological and semantic representations. So, for example, *why didn't he be careful* is not strictly speaking ill-formed; *don't*+V[BARE] is a licensed syntactic form in English. But this example does not satisfy the more idiosyncratic licensing conditions imposed by the constructions WDYB, HDIB and IYDB and so will be judged as unacceptable out of context. However, some examples may be judged as marginally acceptable in context by extension of the broader licensing conditions of the well-established constructions, by assigning an interpretation of directive force and dropping the condition that the subject be the addressee, both conditions on *why don't you be*.

As noted in our discussion of Collins' account of lexical *be* in Section 5, syntactic theories couched entirely in terms of general principles are not capable of accounting for what speakers actually know. Such phenomena provide strong evidence for a non-modular view of grammar in which constructions are ontologically central. On this view, knowledge of language consists of constructions that express correspondences between sound and meaning, mediated by syntactic structure. These constructions range from the very general to the

³¹ Thanks to Satoru Kanno (p.c.) for bringing this possibility to our attention.

very idiosyncratic. The natural state of affairs on such an approach is that grammatical constructions of a language are related to each other, in that more specific constructions acquire properties from more general ones – they take on the shape of non-idiomatic phrases of the same category, or deviate minimally from them (Culicover and Jackendoff 2005; Culicover et al. 2017). We believe that addressing knowledge of language at this level of specificity is essential to ultimately achieving a full understanding of how language works, how it is acquired and how it changes, and how it is represented in the human mind.

References

- Becker, Misha. 2004. *Is isn't be*. *Lingua* 114:399–418.
- Bybee, Joan. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Bybee, Joan. 2013. Usage-based theory and exemplar representations of constructions. In Thomas Hoffmann and Graeme Trousdale (eds.), *The Oxford handbook of construction grammar*, 49–70. Oxford: Oxford University Press.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Collins, Chris. 2006. A fresh look at habitual *be* in AAVE. In Ana Deumert and Stephanie Durrleman (eds.), *Structure and variation in language contact*, 203–224. Amsterdam: Benjamins.
- Culicover, Peter W. 1971. *Syntactic and semantic investigations*. PhD thesis, MIT, Cambridge, MA.
- Culicover, Peter W. 1999. *Syntactic nuts: Hard cases, syntactic theory, and language acquisition*. Oxford: Oxford University Press.
- Culicover, Peter W. 2021. *Language change, variation and universals – a constructional approach*. Oxford: Oxford University Press.
- Culicover, Peter W. & Ray S. Jackendoff. 2005. *Simpler syntax*. Oxford: Oxford University Press.
- Culicover, Peter W., Ray S. Jackendoff & Jenny Audring. 2017. Multiword constructions in the grammar. *Topics in Cognitive Science*. 9. 552–568.
- Culicover, Peter W. & Giuseppe Varaschin. In press. *Deconstructing syntactic theory: A critical review*. Oxford: Oxford University Press.
- Culicover, Peter W., Giuseppe Varaschin & Susanne Winkler. 2022. The radical unacceptability hypothesis: Accounting for unacceptability without universal constraints. *Languages*. 7(2). <https://www.mdpi.com/2226-471X/7/2/96>.
- Denison, David. 2000. Combining English auxiliaries. In Olga Fischer, Anette Rosenbach & Dieter Stein (eds.), *Pathways of change: Grammaticalization in English*, 111–147. Amsterdam: Benjamins.
- Dietrich, Gerhard. 1949. *Die syntax der 'do'-Umschreibung bei 'have', 'be', 'ought' und 'used (to)' auf sprachgeschichtlicher Grundlage dargestellt*. Braunschweig: Westerman.
- Dowty, David. 1979. *Word meaning and Montague grammar*. Dordrecht: Reidel.
- Erades, P. A. 1953. Points of modern English syntax. *English Studies*. 34. 44–46.
- Francis, Hartwell S., Michelle L. Gregory & Laura A. Michaelis. 1999. Are lexical subjects deviant. In *Papers from the 35th regional meeting of the Chicago Linguistic Society*, 85–97. Chicago, IL.
- Goldberg, Adele E. 1995. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Green, Georgia M. 1975. How to get people to do things with words: The whimperative question. In Jerry Morgan and Peter Cole (eds.), *Syntax and semantics 3: Speech acts*, 107–131. New York: Academic Press.
- Hirtle, Walter. 1965. Auxiliaries and voice in English. *Les Langues Modernes*. 59. 25–42.
- Hollenbaugh, Ian. 2023. Habitual *be* in American English. https://ihollenbaugh.files.wordpress.com/2022/07/fs-stowell_hollenbaugh_be-revised-1.pdf. Forthcoming in a Festschrift for Prof. Tim Stowell.
- Huddleston, Rodney. 1976. Some theoretical issues in the description of the English verb. *Lingua*. 40. 331–383.
- Huddleston, Rodney, Geoffrey K. Pullum, et al. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Jackendoff, Ray S. 2002. *Foundations of language: Brain, meaning, grammar, evolution*. Oxford: Oxford University Press.
- Joly, André & Dairene O'Kelly. 1990. *Grammaire systématique de l'Anglais*. Paris: Nathan.

- Kjellmer, Göran. 2009. *How do I be positive?* On factors promoting the DO-support of *be*. *Neuphilologische Mitteilungen*. 110. 73–84.
- Montgomery, Michael & Margaret Mishoe. 1999. He bes took up with a Yankee girl and moved up there to New York: The verb *bes* in the Carolinas and its history. *American Speech*. 74(3). 240–281.
- Partee, Barbara. 1977. John is easy to please. In Antonio Zampolli (ed.), *Linguistic structures processing*, 281–312. Amsterdam: North Holland.
- Payne, Thomas E. 2013. The two *be*'s of English. *Applied Research in Linguistics*. 2. 17–35.
- Potts, Christopher. 2023. Characterizing English preposing in PP constructions. <https://lingbuzz.net/lingbuzz/007495>. Ms., Stanford University.
- Pullum, Geoffrey K. & Barbara C. Scholz. 2001. Model-theoretic and generative-enumerative syntactic frameworks. In Philippe de Groote, Glyn Morrill & Christian Retor'e (eds.), *Logical aspects of computational linguistics*, 17–43. Berlin: Springer.
- Rothstein, Susan. 1999. Fine-grained structure in the eventuality domain: The semantics of predicative adjective phrases and *be*. *Natural Language Semantics*. 7. 347–420.
- Sag, Ivan A. 2012. Sign-based construction grammar: An informal synopsis. In Hans C. Boas and Ivan A. Sag (eds.), *Sign-based construction grammar*, 69–202. Stanford, CA: CSLI Publications.
- Siemund, Peter. 2018. *Speech acts and clause types: English in a cross-linguistic context*. Oxford: Oxford University Press.