



**HAL**  
open science

# On the Apparent Transitivity of Auxiliary Do: A Case Study

Saghie Sharifzadeh

► **To cite this version:**

Saghie Sharifzadeh. On the Apparent Transitivity of Auxiliary Do: A Case Study. *Fuzzy Boundaries in Discourse Studies*, Springer International Publishing, pp.259-285, 2019, 10.1007/978-3-030-27573-0\_12 . hal-03672539

**HAL Id: hal-03672539**

**<https://hal.sorbonne-universite.fr/hal-03672539>**

Submitted on 19 May 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On the Apparent Transitivity of Auxiliary *Do*: a Case Study

Saghie SHARIFZADEH

Sorbonne Université (Paris, France) - English studies

saghie.sharifzadeh@sorbonne-universite.fr

**Abstract.** This paper aims to determine whether the apparent transitivity of the verb *do* in *which* + Noun Phrase (NP) + *do*, *that* + NP + *do* and *this* + NP + *do* constructions can be explained by the presence of a lexical *do* or if *do* is in fact an auxiliary. The electric corpora reveal that *do* is sometimes auxiliary in these structures (e.g. when it refers to a predicate denoting a purely stative event), sometimes lexical (e.g. when it is modified by an adjunct), but that the syntactic properties used as identification criteria for an auxiliary / lexical verb are not always iconic of the semantic-discursive criteria which appeared prototypical of each form of *do* in other structures. One upshot of the semantic unity between lexical and grammatical uses of *do* is the fuzziness of the category boundary between the lexical verb and the auxiliary. By contrasting two corpora, the British National Corpus and the Corpus of Contemporary American English, this study attempts to determine the lexical or grammatical nature of the verb *do* in *prima facie* borderline cases.

**Key words:** *do*, lexical verb, auxiliary, anaphora, transitivity.

## 1. Introduction

This paper investigates *which* NP<sup>1</sup> *do* (cf. (1)), *that* NP *do* (cf. (2)) and *this* NP *do* (cf. (3)) sequences and the apparent transitivity of the verb *do* in its aforementioned uses.<sup>2</sup> The aim is to determine whether such transitivity can be explained by the presence of a lexical *do* in these structures, as in (4), or if *do* is in fact an auxiliary (i.e. a grammatical verb), as in (5). The focus lies solely on the lexical or grammatical status of *do* (considering the possible transitivity of the latter); the subject NPs will therefore be confined to noun phrases headed by personal pronouns (PRPs).

---

<sup>1</sup>

NP = noun phrase.

<sup>2</sup>

- (1) *Perhaps she thinks it sounds n<sup>-3</sup> better. **Which it does** really. Well it does really, yeah.* (BNC KD8)
- (2) *'Would you like my place? It's rather a tight fit,' she said apologetically, and the woman beamed at her. 'Thanks ever so, miss, I do call that kind,' she said breathlessly. 'I wish there was more like you, **that I do!**'* (BNC BMU)
- (3) *I remember one subject that required that she lay flat on the ground, and **this she did** for hours on end while I drew her.* (BNC CN4)
- (4) *Boxing titles are won and lost in the ring and in order to be a champion, you must fight the champion and beat the champion. I **did** that and now (...).* (BNC CBG)
- (5) *I tried to stop you. No you **didn't**. I **did**. No you **didn't**.* (BNC KDE)

After giving a brief summary of the plausible origin of auxiliary *do*, which will prove useful for the following analysis, this paper will scrutinize a certain number of synchronic uses of the verb.

It has long been considered that auxiliary *do* had a causative origin (cf. Abbott 1875; Callaway 1913; Royster 1914–1915, 1918, 1922; Zilling 1918; Ellegård 1953). According to this hypothesis, the semantic shift from the causative (*do somebody* V<sub>INF</sub>) to the periphrastic use (*do* V<sub>INF</sub>) stemmed from metric or prosodic needs in South-Western verse texts. Now, the causative use of the verb appeared as early as the Old English period, and the *Old English Dictionary* (OED) notes that, during this same period, periphrastic *do* – which would later become an auxiliary – appeared for the first time. It coexisted with causative *do* until the end of the Middle English period, before causative *do* progressively disappeared:

General scheme of arrangement – I. Transitive senses (\*To put. \*\*To bestow, render. \*\*\*To perform, effect). II. Intransitive: To put forth action, to act. *III. Causal and Auxiliary uses (\*Causal. \*\*Substitute. \*\*\*Periphrastic).* IV. Special uses of certain parts (Imperative, Infinitive, Pres. pple., Past pple.) (...). (OED, 562: B; emphasis added)

If auxiliary *do* had a causative origin, the causative and periphrastic uses of the verb should not have appeared in the same period of time. Sharifzadeh (2012: 6–135) shows that grammatical *do*, which was originally, almost exclusively, combined with predicates denoting actions or accomplishments (and sometimes activities), is derived from the ‘achieve’, ‘perform’ ... meaning of the lexical verb, cf. (6).

- (6) (?c1450) Wright, T. (ed.) (1906) *The Book of the Knight of La Tour-Landry* rev. edn (EETS, Ordinary series 33), 2–24

*And so thei dede bothe deseieue ladies and gentilwomen, and bere forthe diuerse langages on hem*

and so they did both deceive ladies and gentlewomen and make diverse allegations about them

This lexical *do* progressively grammaticalized into the auxiliary. Hopper et Traugott (1993: XV) define grammaticalisation as “the process whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions”. As an effect of the semantic contribution of the verb, the *do*-construction was at first associated with dynamic events (cf. Kroch, Myhill and Pintzuk 1982; Denison 1993). The auxiliary can nowadays appear with predicates denoting either stative or dynamic events (whether the latter be telic or atelic) but the factitive<sup>4</sup> meaning of *do* is still common in multiple uses of the verb provided that the

---

4

Visser’s acceptance of the term *factitive* will be used throughout the text, to refer to a sense of ‘doing’ or ‘acting’ and not a causative meaning.

latter takes some complementation (a verbal complementation in the case of auxiliary *do*, a nominal complementation – cf. *do it/that/this* – or adverbial complementation – cf. *do so* – in the case of lexical *do*). The etymological background of the verb will even occasionally produce sentences in which, without a thorough analysis of the corpora, it is impossible to determine whether *do* is a lexical verb or an auxiliary. The fuzziness of the category boundary between lexical and grammatical uses of *do* is particularly notable in the structures *which* NP *do* and *that/this* NP *do*, the noun phrases here being limited to PRPs (“[p\*]” in BYU-BNC and BYU-COCA queries).

In the first case, the pro-form (*which*) is a relative pronoun which introduces a non-restrictive relative clause (“supplementary” in Huddleston and Pullum 2002: 1035) and substitutes for a predication; it is followed by a subject PRP and a form of *do* (cf. (1)). In the second case, the pro-form (*that* or *this*) is a demonstrative pronoun functioning as a direct object complement; it is topicalized and followed by a subject PRP and a form of *do* (cf. (2) and (3)).

To identify the properties which allow us to determine if *do* is an auxiliary or a lexical verb, the present study will focus on these two anaphoric structures with a positive *do* (which at face value appears similar to a transitive *do*) + a pronoun (either relative or demonstrative), considering the following questions:

- a. Can *do* be negated (this property being characteristic of operators)<sup>5</sup>?
- b. What types of antecedents does *do* have in such structures? The electronic corpora reveal that purely stative antecedents (e.g. *sound better* in (1)) are extremely rare in pro-forms containing a lexical *do* (cf. Lakoff and Ross 1976; Guimier 1981; Culicover and Jackendoff 2005), particularly with the pro-forms *do it*, *do this* and *do that* (in which, contrarily to *do so*, *do* is transitive). Unlike the lexical *do* of these pro-forms, auxiliary *do* can be anaphoric of predicates referring to dynamic events and stative events equally. The results for *do so* will be revisited later on in the paper as they run somewhat counter

---

5

Cf. Huddleston (1976: 333) and the *NICE* properties (Negation, Inversion, Code, Emphasis).

to most of the opinions expressed in the literature, according to which this pro-form cannot have a stative antecedent.

c. How often is *do* modified by an adjunct? In a study of verbal anaphors with *do*, Miller (2011) tapped into the *Corpus of Contemporary American English* (COCA) to show that adjuncts are rare when an operator governs the anaphora; in a study parallel to that of Miller, Sharifzadeh (2019) established that this is equally the case in the *British National Corpus* (BNC).

This paper will firstly deal with the results relative to the pattern *which* PRP *do* before moving on to an analysis of *that* PRP *do* and *this* PRP *do*. All these structures have been examined in the BYU-BNC as well as in the BYU-COCA.

## 2. *which* PRP *do*

The apparent transitivity of *do* in this type of sequence has led certain authors to classify it as a lexical verb, cf. Méry (2002: 90):

Dans [p. ex. (7)], on analyse sans peine l'occurrence de DO comme étant DOlex, verbe d'action transitif, dont le complément d'objet direct est le pronom relatif *which*, et l'on peut montrer ceci en paraphrasant par *and he did that*.<sup>6</sup>

(7) 'Can you give me a hand?' she asked, *which he did*' (BNC AN7)

Yet the simple fact that *do* can be negated in these sentences (as in (8)) goes against its analysis as a lexical form.

---

6

“In [e.g. (7)], we easily notice the occurrence of DO as a lexical DO, a transitive verb denoting action, whose direct object complement is the relative pronoun *which*, and we can show this by paraphrasing: *and he did that*.”

The translations from French literature are my own throughout the article.

- (8) *And if I didn't follow your car, which I didn't, how on earth would I have known where you lived?* (BNC HGY) (taken from Méry 2002: 90)

Indeed, lexical *do* cannot bear negation:

- (4a) \**I **didn't** that and now (...)*.  
(4b) *I didn't **do** that and now (...)*.

Méry (2002) agrees and proposes the following hypothesis:

Le complément d'objet subsiste, c'est le relatif *which*. On a donc là un type de structure où DO apparaît avec des propriétés mixtes par rapport à celles de DO<sub>lex</sub> et DO<sub>aux</sub>.<sup>7</sup> (2002: 91)

Since sentences such as (1) could be interpreted as in (1a) (which would imply the use of *do* as an operator in these instances), the question arises as to whether – in utterances such as (7) – *do* indeed acts like the lexical *do* present in complex pro-forms<sup>8</sup> (in particular *do that*).

- (1a) *Perhaps she thinks it sounds better. [(And) sound better] it does really.  
Well it does [sound better] really.*

---

7

“The object complement remains, it is the relative *which*. Therefore, we have here a structure where DO appears to have mixed properties in comparison to lexical DO and auxiliary DO.”

8

The term “complex pro-form” is borrowed from Quirk et al. (1985: 866) to refer to pro-forms such as *do it/that/this/so*.

More generally, the aim is to discover whether the possible negation of *do* could be explained by mixed properties, as suggested by Méry (2002).

To obtain finite forms of *do*, the query *which* [p\*] [do] was used to search the BNC, with the objective of retaining only the relative clauses in *which* ... *do* that constitute a subset of those which Quirk et al. (1985: 1118) call “sentential relative clauses”. The possibility of intervening items such as [vm\*] (*which* [p\*] [vm\*] [do]) was excluded, because such queries would necessarily retrieve instances of a lexical *do*; they will therefore be explored separately. According to Quirk et al., each clause from this “sentential relative clause” subset “refers back to the predicate or predication of a clause”; rather, I consider that it is the relative pronoun *which* – not the sentential relative clause as a whole – that has a complete predication as its antecedent (and substitutes for *sound better* in (1)). Other “sentential relatives” are described as referring back to an entire clause, as in “*Things then improved, which surprises me*”<sup>9</sup> (again, I consider that it is *which* – and not the sentential relative in its entirety – that refers to a whole clause). The instances in which the antecedent of *which* was not a predication were eliminated from the data. This is the case in (9) where the relative pronoun substitutes for a noun phrase:

- (9) *A recent book on contemporary public sculpture begins with a sentence the author does not **consider remarkable** but which I do: (...).* (COCA ACAD)

In this statement, there are two anaphoric relationships: *which* substitutes for a noun phrase (underlined in the above), whereas *do* is followed by the post-auxiliary ellipsis (PAE) of a verbal antecedent (in bold). *Which I do* thus means ‘[a sentence] I do [consider remarkable]’, not \*[consider a sentence remarkable] I do’: the relative clause does not fulfil the requirement of containing a *which* that refers back to a whole predication. These cases are therefore irrelevant to the context of the present analysis. The focus will be on utterances

---

9

Underlining added.

Throughout this study, the antecedent will be underlined.



such as (1) or (7), to attempt to determine whether the structure *which ... do* should be read as:

- a relative pronoun followed by an operator, or
- a relative pronoun followed by a lexical *do* in a type of anaphora comparable to substitutions by a transitive *do* + an object complement (*do that* in particular), or even
- a relative pronoun followed by a mixed form of *do*.

## 2.1. The finite *do*

### 2.1.1. Results in the BNC

1181 hits of the sequence *which* + [p\*] + a form of *do* (negated or not) were obtained, from which a random sample of 200 was taken. Of this sample, 56 occurrences were relevant. By extrapolation, it can be estimated that there are about 330 occurrences of this pattern in the BNC, so 3.3 instances per million words<sup>10</sup> (cf. Table 1).

| Sequence               | PER MIL                 | polarity <i>do</i> % <sup>‡</sup> |      | antecedent % <sup>‡</sup> |         |       | adjunct % <sup>‡</sup> |
|------------------------|-------------------------|-----------------------------------|------|---------------------------|---------|-------|------------------------|
|                        |                         | +                                 | -    | dynamic                   | stative | mixed |                        |
| <i>which</i> [p*] [do] | 3.31 (#56) <sup>†</sup> | 80.4                              | 19.6 | 89.3                      | 10.7    | -     | 61                     |

**Table 1** Properties of the structure corresponding to the sequence *which* [p\*] [do] (*n't/not*) (with a finite *do*) in the BNC

<sup>†</sup> Given that the query excludes the possibility for an adverb (or more) to be interpolated between the subject pronoun and the form of *do*, this does not represent an exhaustive number.

<sup>‡</sup> These percentages are relative to the number of occurrences surveyed (or extrapolated) for each sequence (here of a total of 56 instances).

In the shaded areas of the tables, the number of occurrences examined corresponds to a total obtained from samples ranging from 100 to 2000 occurrences (here 200). In each table, the number of occurrences observed has been noted in brackets, so that the reader keeps in mind that the proportions are often expressed on the basis of restricted numbers of occurrences.

The number per million gives a perfect idea of the frequency in the corpus. In order to synthesize the information, square brackets are used – as in the BYU interfaces – to refer to all word-forms of the lemma *do*, which are necessarily finite in this distribution. It is specified in the legend that they can be negated.

Of the sample in Table 1, 11 negative occurrences (19.6%) were identified (e.g. (10) and (11)). Regarding the type of event referred to by the antecedent, 6 cases (10.7%) are stative, versus 50 dynamic (89.3%).<sup>11</sup> It should be noted that the stative antecedents include purely stative ones, that is, events denoted by a predicate whose subject cannot be understood as agentive, whether in the antecedent or in the anaphoric clause (cf. examples (1) and (11)).

(10) *This is not because of Finn making jokes (**which he doesn't**), it is from things which actually happen to him (...)* (BNC KA1)

(11) *Not unless they had radar or heat-seeking missiles, **which they didn't**. Did they?* (BNC HWL)

These two possibilities (negative form and stative antecedent) indicate a close proximity between the verb *do* in *which* PRP *do* structures and the auxiliary (e.g. *No, you didn't*) which, unlike lexical *do*, can bear the negation and equally have stative or dynamic antecedents. These observations also bring to evidence a marked difference from complex pro-forms, particularly *do that/this*, in which it is clear that *do* is a lexical verb which takes the demonstrative pronoun as a direct object complement (*\*He didn't that/this*).

---

<sup>11</sup>

See Sharifzadeh (2019) for the method adopted in determining the lexical aspect (stative/dynamic, telic/atelic) of the antecedent.

In his comparative study of PAE and complex pro-forms, Miller (2011: 7) made the following observation based on the COCA data:

Another important factor affecting the choice of PAE is the presence of an adjunct after the elliptical auxiliary. Though this is not ungrammatical, it is rare. Out of 249 cases of PAE examined, only 4 exhibited this pattern, i.e. 1.6%.

Miller infers that the presence of an adjunct decreases the likelihood of PAE. In the rare cases observed (e.g. *Brother Laurence has rarely grasped for words. He does now.*), the author considers that what makes the PAE conclusive is that:

- (i) there is a clear polar alternative between the habitual situation and the present / specific situation
- (ii) the adjunct does not characterize the situation referred to by the antecedent, as is the case in all the examples of *do so* with adjuncts cited [earlier in the paper], but rather contrasts the specific situation referred to by [*do*] with the habitual situation referred to by the antecedent. (2011: 12)

On the basis of these observations, it was deemed useful to determine the frequency with which *do* is accompanied by an adjunct in those relative clauses introduced by *which*, given that the presence of adjuncts might constitute an additional criterion for determining whether, in such anaphoric contexts, *do* exhibits features that are similar to those of the auxiliary or those of the lexical verb found in complex pro-forms.

Contrary to the above observations, which were based on the existence of negative forms and the possibility of a stative antecedent and which therefore implied the use of a purely grammatical *do* in such contexts, there is a marked difference between the verb *do* in these instances and the operator in cases of PAE. More than half of the occurrences of *do* are accompanied by an adjunct; yet, notably, the operator in PAE and the anaphoric non-finite *do* (in cases such as *yes/no* + PRP + auxiliary + *do*, e.g. *Yes, I should do*) both tend to appear without an adjunct. From this point of view, the operator and the non-finite substitute *do* differ from *do so* and *do it/that/this* (52% of the *do so* and *do it/that/this* occurrences

examined in Sharifzadeh (2019) appear with an adjunct). The following is an example of the sequence *which* + [p\*] +[do] with two adjuncts:

- (12) [*He was*] *the only Briton to win a world Junior Title, which he did in Florida in the 14 and under age category.* (BNC CKL)

The adjuncts ‘*in Florida*’ and ‘*in the 14 and under age category*’ carry additional information about the event mentioned in the antecedent, which is also the case with *do so*. After a close examination of all occurrences, it appears that only two instances (which will be revisited further on in the paper) consist of a negated *do* accompanied by an adjunct, in a situation where the criteria (i) and (ii) (described above) are met:

- (13) *If we are to believe late nineteenth- and early twentieth-century etiquette writers (**which I don't** altogether) this was a period when it was thought gross to (...).* (BNC EFU)
- (14) *I normally get that first thing in the morning, **which I didn't** today. So you'll have to make do with scones.* (BNC AN7)

These results are not overly surprising given the semantic contribution of the clause containing *which*. When this clause has a positive polarity, it serves to fulfil two functions. The first function is relative to cases in which the event denoted by the antecedent is stated or presupposed; the *which* clause then marks a confirmation of the validity of the event (cf. (12)). However, this confirmation would often be completely redundant were it not accompanied by additional information, which is provided by the adjunct:

- (12a) #*He was the only Briton to win a world Junior Title, which he did.*

For discursive reasons, these cases here, which are parallel with the majority of *do so* occurrences (e.g. *They judged that Mr Mandela and President de Klerk had laid the foundations for a new democratic South Africa and **done so** through personal integrity and great political courage* (BNC K6C)), almost impose the presence of an adjunct.

The second function of the relative clause containing *which* is shown below:

- (15) *Our orders were to keep a low profile while keeping watch on the yacht **which** we **did** for about twenty four hours. (BNC H0C)*

The event denoted by the antecedent has not occurred; it is presented as potential. Unsurprisingly, all relatives of a positive polarity without an adjunct refer back to an antecedent that presents an event as potential. By employing a *which*-clause, the speaker informs us that this potential event has been completed. This second function also appears with *do so*; the following is an example of this:

- (16) *The Government wanted to create a property-owning democracy, and they **did so**, but that property-owning democracy needs inflation to maintain itself. (BNC HHW)*

*Do so*, like *which* [p\*] [do], allows the speaker to assert the accomplishment of an event, stated as potential in the antecedent clause. When the *which*-clause has a negative polarity, it serves to deny the accomplishment of the event denoted by the antecedent (be it actual or potential). In the rare cases where a negated *do* is followed by a complement (cf. (13) and (14) above), the latter serves to indicate that the event is only partially invalidated. In (13), the speaker indicates that his disbelief is not total and, in (14), he indicates that the accomplishment of the event (namely, receiving the patisserie in question first thing in the morning) is only invalid on that particular day, and remains valid the rest of the time (*normally*).

After examining the characteristics of the structure *which* + PRP + finite *do* in the BNC, a partial resemblance can be noticed to cases of PAE with *do* (cf. the possibility of a stative antecedent, whatever it may be, and a negative form) mixed with a partial resemblance to *do so* (cf. the types of anaphora, the frequency of adjuncts, the functions of the latter). This configuration thus seems to provide evidence for an area of interpenetration between the grammatical and the lexical, where certain traits of the lexical spill over into the grammatical. This is perhaps due to the apparent transitivity of *do* in these subordinate relative clauses, a transitivity which is specific to the lexical verb and which has led Méry (2002) to liken these uses to the pro-form *do that*. There is no similarity with *do that*, whose properties are very different from those found in these clauses: this complex pro-form always refers to a dynamic

event and cannot be negated. The electronic corpora reveal that *do so* has the ability to substitute for purely stative antecedents<sup>12</sup> – albeit more rarely than the clauses which interest us here – but, although it bears the factitive meaning characteristic of the lexical, *do* is not transitive and here means ‘act’ rather than ‘accomplish’. It is, therefore, the structure *which ... do* that is comparable to the pro-form *do so*, and not the verb *do* taken in isolation in each case. The syntactic test remains nevertheless problematical in terms of the assimilation, in these clauses, of the verb *do* with a lexical verb: if *do* in the structures *which [p\*] [do]* were lexical, it could not be negated. More precisely, the negative counterpart of *which [p\*] [do]* should then be *which [p\*] [do] n t/not do* (the first *do*, an operator, acting as support for the second, lexical). The frequency of the sequence *which [p\*] [do] n t/not do* was verified and only 2 instances were found in the entire BNC. For the record, of the 56 relevant occurrences of *which [p\*] [do]* (positive and negative) which were listed without a verb following the auxiliary, 11 were *which [p\*] [do] n t/not* cases. In other words, in nearly 20% of these occurrences, *do* was in its negative form (cf. the NICE properties of operators). Such a gap between the queries *which [p\*] [do] n t/not do* and *which [p\*] [do] n t/not* in British English proves that, on a syntactic level, *do* is indeed an operator in this configuration, even if it appropriates certain semantic-syntactic properties of the lexical. Can the same be said for American English?

### 2.1.2. Results in the COCA

80 relevant occurrences of the sequence *which [p\*] [do]* were taken from a random sample of 200 (out of a total of 3821). By extrapolation, this is equivalent to 1530 in the whole

---

<sup>12</sup>

E.g.:

*Though something in the provenance or context of the document may justify Husameddin's assertion, it must be noted that the signature itself does not **do so**.* (BNC H7S)

*'An old flame, perhaps,' suggested Dorothea who, often an unthinking woman, felt vaguely that the afternoon might now benefit from a stroke of the unlikely. She should not have **done so**.* (BNC AD1)

corpus, hence 3.6 per million words. This frequency is very close to that found in the BNC (see Table 2).

Among the occurrences examined, there were negative ones (13.8%) as well as relatively numerous cases in which the antecedent denotes a stative event (17.5%). Once again, the predicates concerned may refer to a purely stative event, e.g.:

(17) *Even if he had known the language, **which he didn't**, of course, Peterson would have had trouble asking the conductor (...). (COCA FIC)*

As in the BNC, *do* is often accompanied by an adjunct (46% of cases) and the following two types of positive-polarity clauses can be found:

- Actual event in the antecedent, e.g.:

(18) *This is an intellectual swindle that leads women to misjudge male sexuality, **which they do** at their own emotional and physical peril. (COCA MAG)*

- Potential event in the antecedent, e.g.:

(19) *If TGF-inhibitors (...) worked efficiently, **which they do**, its levels would plummet below a certain threshold. (COCA MAG)*

Regarding the 11 negated occurrences, none of them are modified by an adjunct. As was suggested above, there are no instances of an adjunct because the event denoted by the antecedent is simply denied. There could, however, have been more occurrences of a partial invalidation of the event referred to in the antecedent (cf. (13) and (14)), as is the case with more than half of the instances of the sequence *which [p\*] [do] n't/not do* in the COCA (see below, e.g. *'I save everything she sends me, which I don't do with anyone else,' he said. (COCA)*).

This difference in proportion can be explained by the fact that, in the pattern *which [p\*] [do] n't/not*, the fuzziness of the boundary between the lexical and the grammatical is syntactically blocked by the negation, which can only be borne by the auxiliary (the *N* property being reserved to operators). *Do* could therefore retain more of the characteristics

which are specific to the operator in other contexts (notably the absence of an adjunct, typical in anaphoric utterances with a grammatical *do*).

Table 2 summarizes the results which have just been discussed.

|                            | Sequence               | PER MIL    | polarity <i>do</i> % |      | antecedent % |         |       | adjunct % |
|----------------------------|------------------------|------------|----------------------|------|--------------|---------|-------|-----------|
|                            |                        |            | +                    | -    | dynamic      | stative | mixed |           |
| <b>B<br/>N<br/>C</b>       | <i>which</i> [p*] [do] | 3.31 (#56) | 80.4                 | 19.6 | 89.3         | 10.7    | -     | 61        |
| <b>C<br/>O<br/>C<br/>A</b> | <i>which</i> [p*] [do] | 3.6 (#80)  | 86.2                 | 13.8 | 82.5         | 17.5    | -     | 46        |

**Table 2** Properties of the structure *which* [p\*] [do] with a finite *do* (negated or not) in the BNC and in the COCA

Given these results, it might be tempting to conclude that the properties of *do* in this configuration are the same in British and American English. However, there is a notable difference between the two corpora when it comes to *which* [p\*] [do] *n't/not do* occurrences: of 55 occurrences retrieved by the query in the COCA, 38 are negations of sentential relatives that meet our selection criteria. In other words, American English could more easily apply the syntactic properties of the lexical verb to *do* in these clauses (2 occurrences in 100 million words in the BNC / 38 occurrences in 450 million words in the COCA). Given that this difference in proportion may not be sufficiently significant, the next step was to verify whether the properties of non-finite *do* in the pattern *which* [p\*] [do] *n't/not do* (cf. (20)) were the same as those of finite *do* in the pattern *which* [p\*] [do].

- (20) *Puzzled, he wondered if she had gone upstairs to her children, **which she did not do** as a rule, and was about to ascend; when from up there came a great cry, and the sound of a fall from the window.* (BNC CD2)

## 2.2. The non-finite *do*

Faced with utterances similar to (20), which highlight the non-auxiliary status of *do* in the anaphora, it was deemed necessary to explore *which ... do* structures with a non-finite *do*, which cannot bear the negation itself. The *N* property will be borne by another auxiliary, e.g.



auxiliary *do*, as in (20) (*did not*), or a modal auxiliary (“[vm\*]” in the BYU interface). On a strictly syntactic level, if *do* is non-finite it must be lexical in this pattern, for it is the unit *which ... non-finite do* which substitutes for the predicate (non-finite *do* cannot be periphrastic, as is shown in (20a) and (20b)).

(20a) *Puzzled, he wondered if she had gone upstairs to her children. \* She did not do [go upstairs to her children] as a rule.*

(20b) *\*He wondered if she had gone upstairs to her children, [(and) go upstairs to her children] she did not do.*

It is still important to consider the polarity when discussing the presence / absence of an adjunct in these clauses with a non-finite *do* (cf. Table 3). The possibility / impossibility of encountering stative antecedents will also be investigated, in order to determine if *do* is purely lexical here or if it is lexical with (partial) grammatical characteristics.

|      | Sequence                   | PER MIL                   | polarity <i>do</i> % |      | antecedent % |                  |       | adjunct % |
|------|----------------------------|---------------------------|----------------------|------|--------------|------------------|-------|-----------|
|      |                            |                           | +                    | -    | dynamic      | stative          | mixed |           |
| BNC  | <i>which [p*] [do] do</i>  | 0.07 (#7)                 | 71.4                 | 28.6 | 100          | 0                | -     | 57        |
|      | <i>which [p*] [vm*] do</i> | 0.5 (#50)                 | 71.4                 | 28.6 | 91.8         | 8.2              | -     | 65        |
|      | <i>which [p*] [vm*]</i>    | 0.4 (#6)                  | 66.7                 | 33.3 | 100          | 0                | -     | 17        |
| COCA | <i>which [p*] [do] do</i>  | 0.09 (#40)                | 5.1                  | 94.9 | 97.4         | 2.6              | -     | 51        |
|      | <i>which [p*] [vm*] do</i> | 0.5 (#62+64) <sup>†</sup> | 69.8                 | 30.2 | 97           | 3                |       | 72        |
|      | <i>which [p*] [vm*]</i>    | 0.38 (#13)                | 30.8                 | 69.2 | 92.3         | 7.7 <sup>‡</sup> | -     | 23        |

**Table 3** Properties of the structures *which [p\*] [do]/[vm\*] do* (i.e. with non-finite *do*) and *which [p\*] [vm\*]*, in the BNC and the COCA

<sup>†</sup> Since the data for positive occurrences were taken from a random sample of 100 of the 239 responses to the positive query, all the figures for POSITIVE must be multiplied by 2.39 before being added to the figures for NEGATIVE. Of the 100, 62 were relevant. To obtain an accurate calculation (i) of the total number of occurrences of the sequence (i.e. positive cases + negative cases) and (ii) of the POSITIVE / NEGATIVE distribution in the corpus, it was necessary to:  
 - multiply 62 by 2.39 = 148,  
 - add the 148 positive occurrences to the 64 negative occurrences, which gives a total of 212 occurrences of the sequence in the corpus (69.8% positive vs 30.2% negative). It was therefore necessary, for the sequence *which [p\*] [vm\*] do* in the CoCA, to divide by 212.

<sup>‡</sup> This refers to one single occurrence:

*In any case, they've gone off somewhere together. To celebrate before the rest of the world knows anything, **which** they will, soon, from coast to coast and in several large foreign cable markets – and that's as close to a direct quote as I think I can get.* (CoCA)

As can be seen in rows 1, 2, 4 and 5 of Table 3, the *which*-relative clauses contain an adjunct in more than half of the cases. This result is similar to those obtained for finite *do* in the sequence *which* [p\*] [do] and in complex pro-forms.

This situation may be explained (at least in part) by the discursive functions fulfilled by *which*-clauses, at least those with positive polarity. As was previously indicated, these discursive functions encourage the appearance of an adjunct. The *which*-clause in (21) does not require the presence of an adjunct not to be redundant, because the antecedent refers to a potential event.

- (21) *I believe that, you know, in the last six months or eight months she wanted to take a lot of time off to write her book, which she did do, (...).* (COCA NEWS)

Regarding the types of antecedent, the figures confirm that *do* is lexical in these non-finite uses: for the sequence *which* [p\*] [do] *do*, the BNC does not provide a single occurrence with a stative antecedent, and the COCA provides only one, attributed to the English actress Joan Collins. The antecedent is, at first sight in any case, a prototypical stative:

- (22) *And she also bears grudges, which is another thing that I don't do. And she has huge hatreds which I don't do. I don't have hatreds.* (COCA SPOK)

This statement can probably be understood in two different ways. In the first, the predicate *<have huge hatreds>* is understood as the antecedent of *which*. In this interpretation (which can be paraphrased by *having huge hatreds is something I don't do*), the antecedent is evidently stative. Given that this is an isolated case, the analysis of the structure “non-finite *do* + object pronoun *which*” as a lexical *do* + a pronoun, substituting for the predicate *have huge hatreds*, cannot be dismissed. However, a second interpretation is possible, which would make the noun phrase *huge hatreds* the antecedent of *which*; the clause would then mean *I don't do huge hatreds*, with a general *do* similar to the use underlined in (23):

- (23) *I don't do hatred.*

*I really don't. Not ever. What would be the point? Life is long, people are foolish, the world continues regardless. Why waste energy on fury, futile resentment, hate?* (Charlotte E. English, *Evastany, The Draykon Series, Book 5*, <http://www.charlotte-english.com/book/evastany>)

It seems difficult to pick one interpretation over the other without falling into arbitrariness but, even if one opts for the first interpretation, it can be argued that the anaphor itself refers to a dynamic event. This is the case in the few occurrences of a complex pro-form such as *do it/that/this* with a stative antecedent (cf. Sharifzadeh 2019), e.g.:

- (24) (...) *our friendly architect informed us that the Secretary of State was opening the building four weeks earlier than scheduled and it would be very nice to have the furniture there. The answer to such situations is to always say 'Yes, sir, of course we can **do it**,' and then put the telephone down, swear loudly and sit down and work how you are going to do it.* (BNC EFH)

The substitute *do it* in (24), anaphorically linked to *have the furniture there*, does not have the same meaning as its antecedent. The anaphor here is not to be understood as *have the furniture there*, rather as *get the furniture there*; the event denoted by the pro-form is therefore dynamic. Following the first interpretation in (23), Joan Collins would indicate that she (deliberately) chooses to not maintain excessive hatred, rather than informing us that she doesn't possess this trait – experiencing enormous hatred – irrespective of her will.

While, on both a syntactic and semantic level, the lexical quality of *do* in the sequence *which* [p\*] [do] *do* is incontrovertible, can the same be said for the other non-finite variants of *which* structures? The occurrences of the sequence *which* [p\*] [vm\*] *do* were also measured across the two corpora. In 50 relevant occurrences in the BNC, 4 have a stative antecedent.

The distributions were actually measured based on only 49 occurrences, as two of them were identical. Of the 126 relevant occurrences in the COCA, 5 have a stative antecedent. These 9 stative occurrences are listed below:

- (25) *and er we we would like to see a report back **which** we will **do** anyway erm (...).* (BNC JS7)

- (26) *I'm going to see what it does when it's stuck onto carbon, **which** it will **do** very readily. (BNC KRG)*
- (27) *(...) family should be warned against condoning the patient's anorexic behaviour; **which** they may **do** either because they feel unable to tolerate the patient's anger or because (...). (BNC HWW)*
- (28) *Their eyes are very elaborate. In some ways they are even better than our own, for a squid can distinguish polarised light **which** we cannot **do** and their retinas (...). (BNC EFR)*
- (29) *She has to be sexy, **which** she can **do**. (COCA NEWS)*
- (30) *(...) and all of a sudden you hear the countermelody, **which** you wouldn't **do** if you were sitting in the back of the house. (COCA NEWS)*
- (31) *(...) provided they adhere to the laws of war, **which** they will not **do** in any event. (COCA ACAD)*
- (32) *Archaeology perhaps seems “old” because we associate it with early civilizations (**which** we need not **do**), (...). (COCA ACAD)*
- (33) *Such an instrument (...) lets you see in depth, **which** you can't **do** very well with one eye. (COCA MAG)*

What is more striking than the (low, especially in the COCA) number of cases which have a stative antecedent is the type of state predicates which can be seen – many of which refer to “purely” stative events. Even more striking is the fact that in several cases there seems to be absolutely no justification for interpreting the anaphor as dynamic. In all likelihood, (29) and (31) – and possibly (27) and (33) – lend themselves to such an interpretation, but in the other cases, the relative clause is naturally interpreted as stative. This observation differentiates *which...* non-finite *do* from *do it/that/this*. However, the anaphora is similar to cases of PAE with *do* as an operator but also to substitutions with *do so*. Since a high frequency of adjuncts

was found, the strongest comparison seems to be with *do so*, even though the operator itself may assume the properties of *do so* in *which*-clauses, which is particularly interesting to us here.

The factors that were put forward in the present study as identification criteria for an auxiliary (based on criteria also used by other authors, notably Miller 2011) do not seem to appear systematically in all syntactical environments. Consequently, a pattern associated with a low number of stative antecedents but with a high number of adjuncts would not necessarily be classified among the structures containing a lexical *do*. In certain contexts, the syntactic criteria are no longer iconic of the semantic-discursive criteria which appeared prototypical of each form of *do*: in the case of the non-finite *do* present in *which ... do* clauses, the properties of the grammatical spill over into the lexical, just as the lexical spilled over into the grammatical in the finite form.

This structure highlights an occasional difficulty in differentiating between the lexical and the grammatical. Is the same true for *that/this* [p\*] [do] clauses, in which *that/this* can be read, at first glance, as a potential object of the verb *do*?

### 3. *that/this* PRP *do*

Based on the model of Mery's (2002) analysis of *which* PRP *do* substitutions, the structure *that/this* PRP *do* would be analysed as resulting from the topicalisation of the demonstrative pronoun from an initial substitution with *do that*:

(7) 'Can you give me a hand?' she asked, *which* he did. (BNC AN7)

(7a) ?? 'Can you give me a hand?' she asked, and he did *that*.

(7b) 'Can you give me a hand?' she asked, and *that* he did.

Like *which ... do* anaphors, *that/this ... do* would in this case have the semantic and/or discursive properties of the pro-forms *do that/this*, even if *do* in such instances possesses the syntactic properties of the auxiliary, cf. the *N* property in (34):

(34) *I don't know why I put up with you and your arrogant ways, **that** I don't!*  
(BNC AN8)

A close examination of the electronic corpora will allow us to determine if, in this pattern, *do* acts like an operator or like the verb *do* in complex pro-forms. The left context had to be limited in the corpus query, so as to avoid excessive numbers of irrelevant occurrences linked to the presence of *that*. The two queries were as follows:

*and/but/though that/this* [p\*] [do]

*./,;/;/?!/: that /this* [p\*] [do]

### 3.1. Finite *do*

Table 4 provides a summary of the results. As in the previous tables, the percentages given are based on the total number of occurrences surveyed (or extrapolated) for each sequence; the cells which contain figures derived from extrapolations of random samples have been shaded. Here we are dealing with 2 samples of 200 occurrences from 1095 for the left context *and/but/though*, and 2 samples of 200 occurrences from 2898 for the left context *./,;/;/?!/:*.

|                            | Sequence                | PER MIL    | polarity <i>do</i> % |    | antecedent % |         |       | adjunct % |
|----------------------------|-------------------------|------------|----------------------|----|--------------|---------|-------|-----------|
|                            |                         |            | +                    | -  | dynamic      | stative | mixed |           |
| <b>B<br/>N<br/>C</b>       | X <i>that</i> [p*] [do] | 0.07 (#7)  | 71                   | 29 | 43           | 57      | -     | 14        |
|                            | X <i>this</i> [p*] [do] | 0.84 (#84) | 100                  | 0  | 95           | 4       | 1     | 71        |
| <b>C<br/>O<br/>C<br/>A</b> | X <i>that</i> [p*] [do] | 0.21 (#19) | 89                   | 11 | 83           | 11      | 6     | 33        |
|                            | X <i>this</i> [p*] [do] | 0.18 (#75) | 100                  | 0  | 99           | 1       | -     | 85        |

**Table 4** X *that/this* [p\*] [do] in the BNC and the COCA  
(X represents conjunctions or punctuations)

Let us first consider the results for *that* [p\*] [do].

#### 3.1.1. *that* PRP *do*

- Results in the BNC

In total, for the two queries, 585 occurrences were obtained (189 + 396) and these were examined one by one. Of these, only seven proved to be relevant. Of this very small (but exhaustive for the two selected left contexts) number, it appeared that two occurrences were

negative and four had a stative antecedent (*want you begging back here; wish there was more like you; know of the work; know why I put up with your arrogant ways*). One occurrence had an adjunct.

Given the possibility that *do* bears negation, combined with the possibility of referring back to a stative antecedent and the relative rarity of *do* being modified by an adjunct, the verb *do* in this structure appears to be an operator. To confirm this observation in American English, the same queries were carried out in the COCA. The results are mixed.

- Results in the COCA

The number of occurrences is not exhaustive with a topicalized *that* in the COCA. 400 occurrences within the constraints of the left context were examined – of 1095 for *and/but/though that* [p\*] [do] and of 2898 for *./,;/;/?/!/: that* [p\*] [do]. The frequency PER MIL in Table 4 is therefore the result of an extrapolation that was not necessary for the other data in the table. Of the 18 relevant occurrences observed<sup>13</sup>, 15 have a dynamic antecedent, one has an antecedent between stative and dynamic and two have a clearly stative antecedent (e.g. (35)).

(35) *By definition, Photo-Travel should have world-wide variety, and **that** it **does**.* (COCA ACAD)

Adjuncts are relatively frequent (one third of cases). There is one negative occurrence, in a novel, which – it should be noted – is supposed to be produced by a Scotsman and not by an American:

(36) *“I suppose I needn't tell a Scotsman about English wars and English taxes.” #  
“Aye, **that** you **don't**,” I agreed.* (COCA FIC)

---

13

There were in reality 19, but only 18 were counted as one of them appears in two samples.

In all, the evidence in favour of an auxiliary *do* in this structure seems less conclusive in the COCA than in the BNC. It could be, then, that this *do*, in American English, is a little closer to the lexical *do* than it is in British English. Nevertheless, we must remain extremely cautious given the very low number of occurrences on which these conjectures are based in the COCA. Let us now turn to the sequence *this PRP do*.

### 3.1.2. *this PRP do*

Firstly, it is noteworthy that, in the BNC, but not in the COCA, the occurrences are much more numerous than *that PRP do* occurrences: 84 versus 7 in the BNC, but 75 versus 88 (with extrapolation) in the COCA. The frequency of the pattern is therefore much higher in the BNC, since COCA is 4.5 times larger than the BNC in terms of word tokens.

This time, the results are extremely coherent between the BNC and the COCA and, contrary to expectation, they are also very different to those obtained for *that*. In summary,

- there were no negated occurrences;
- the type of event denoted by the antecedent is almost systematically dynamic (potential exceptions are discussed below);
- in the vast majority of cases, *do* is modified by an adjunct.

The conclusion seems clear: in the sequence *this [p\*] [do]*, *do* is without doubt a lexical verb. Its profile is very similar to that of the finite *do* in complex pro-forms, particularly *do so*, as is shown in the examination of the few stative cases<sup>14</sup>:

(3) *I remember one subject that required that she lay flat on the ground, and **this** she **did** for hours on end while I drew her.* (BNC CN4)

(37) *The result is that Dublin has to stand on its own constant, as well as temporary, merits. **This** it **does** admirably.* (BNC G21)

---

14

Example (3) is repeated here for the convenience of the reader.



- (38) *Golding's novel is to some degree experimental in style, and when we read it, we sense there is something "odd" about Lok's language. **This we do** by reference to primary norms. (BNC EWA)*
- (39) *The whole operation was supervised from the 1780s from Goodwood by the Duke of Richmond. A general in the regular army, he could only serve as lieutenant-colonel of the Sussex militia. **This he did** as a public duty until the early nineteenth century (...). (BNC CB6)*
- (40) *Dr. Hayes is hardly a rounded character in the novel. We don't learn much about him, he is simply there to represent modern neurology, and **this he does** very well. He's the man with the answers. (COCA ACAD)*

Even if the anaphor in (3) clearly refers to a stative event (“*for hours on end*”), the antecedent could have been interpreted as both stative (as evidenced by logical tests, e.g. *When I came back into the room, she **still** lay flat on the ground*) and inchoative (*When I came back into the room, she lay flat on the ground (=she proceeded to lie flat on the ground)*). The antecedent in (37) can only be interpreted as stative: *Dublin stands on its own merits* may refer to an event which is valid at the time of speech; the test of adding a subordinate temporal clause is also conclusive: *When the local council launched the renovation plan, Dublin **already** stood on its own merits*.

The antecedent in (38) equally describes a state. The verb *sense* belongs to the same category as verbs denoting mental or perceptual states such as *think*, *feel* or *see*.

The antecedent in (39) also refers to a stative event: *When I was appointed governor-general, he **already** served as lieutenant-colonel of the Sussex militia*.

The antecedent in (40) would also be interpreted as stative: *When I first met Mike, he **already** represented modern neurology*. Given the meaning of *represent* here, it seems impossible to interpret this as dynamic.

Consequently, examples (37), (38), (39) and (40) have been classed as stative. (3) has been placed in a different subsection, under the heading “mixed” in Table 4. These proportions, as well as the high frequency of adjuncts, almost mirror the functioning of the pro-form *do so*

(rather than the pro-forms *do it/that/this*, which are equally compatible with adjuncts but are incompatible with a stative interpretation of the anaphors).

### 3.2. Non-finite *do*

To obtain a more complete picture of the sequences *that* [p\*] ... non-finite *do* and *this* [p\*] ... non-finite *do*, a selection of queries was conducted in the BNC and in the COCA, by imposing, as was done for the finite *do*, either a conjunction (*and/but/though*) or a punctuation mark (*./;/;/?/!/:*) in the left context.

|                  | Sequence                           | PER MIL    | polarity <i>do</i> % |     | antecedent % |         |       | adjunct % |
|------------------|------------------------------------|------------|----------------------|-----|--------------|---------|-------|-----------|
|                  |                                    |            | +                    | -   | dynamic      | stative | mixed |           |
| B<br>N<br>C      | <i>X that</i> [p*] [do] <i>do</i>  | 0.01 (#1)  | 0                    | 100 | 100          | 0       | -     | 0         |
|                  | <i>X this</i> [p*] [do] <i>do</i>  | 0.03 (#3)  | 0                    | 100 | 100          | 0       | -     | 0         |
|                  | <i>X that</i> [p*] [vm*] <i>do</i> | 0.07 (#7)  | 57                   | 43  | 100          | 0       | -     | 29        |
|                  | <i>X this</i> [p*] [vm*] <i>do</i> | 0.13 (#13) | 77                   | 23  | 100          | 0       | -     | 85        |
|                  | <i>X that</i> [p*] [vm*]           | 0.04 (#1)  | 0                    | 100 | 100          | 0       | -     | 0         |
|                  | <i>X this</i> [p*] [vm*]           | 0          | -                    | -   | -            | -       | -     | -         |
| C<br>o<br>C<br>A | <i>X that</i> [p*] [do] <i>do</i>  | 0.01 (#6)  | 33                   | 67  | 100          | 0       | -     | 0         |
|                  | <i>X this</i> [p*] [do] <i>do</i>  | 0.01 (#3)  | 33                   | 67  | 100          | 0       | -     | 33        |
|                  | <i>X that</i> [p*] [vm*] <i>do</i> | 0.04 (#17) | 29                   | 71  | 100          | 0       | -     | 12        |
|                  | <i>X this</i> [p*] [vm*] <i>do</i> | 0.05 (#23) | 83                   | 17  | 96           | 4       | -     | 43        |
|                  | <i>X that</i> [p*] [vm*]           | 0.06 (#1)  | 100                  | 0   | 100          | 0       | -     | 0         |
|                  | <i>X this</i> [p*] [vm*]           | 0.002 (#1) | 0                    | 100 | 100          | 0       | -     | 0         |

**Table 5** *X that/this* [p\*] [do]/[vm\*] *do* in the BNC and the COCA  
(X represents conjunctions or punctuations)

Many of the remarks made at the beginning of section 2.2. for *which* [p\*]... non-finite *do* equally apply here. In particular, it is not clear if polarity is a useful indication, since the negation – if there is negation – would be carried by an auxiliary different from the verb *do* which is of interest to this study; the polarity does, however, explain the complete lack of adjuncts in “*that/this* ... [do] *n*’*t/not do*” clauses, simply intended to invalidate the event denoted by the antecedent. On the other hand, the fact that the antecedent is almost systematically dynamic (for *that* as well as for *this*) should certainly be highlighted. Of all of the examples examined, only one has a static antecedent:

- (41) *And him, he may draw you against him, asking you to be brazen and trembling, mutable and firm. **This** you can **do**. You are the sight of you, the touch of you, your scent your scent and sound and will (...). (COCA FIC)*

This occurrence is similar to the data found on complex pro-forms, particularly *do that/this*: the antecedent is indeed stative but it could reasonably be argued that the anaphor refers to a dynamic event (or at least to an event which is deliberately accomplished by an agent). On the basis of these data, it can be hypothesized that instances of *that/this ... non-finite do* are comparable to *do that/this* substitutions with a topicalized demonstrative pronoun.

#### 4. Conclusion

The analysis of the data concerning the sequences *which PRP do*, *that PRP do* and *this PRP do* made it possible to see to what extent the transitivity of the lexical *do* is reflected in some of the auxiliary uses of the verb. These structures reveal the fuzzy boundary between the lexical and the grammatical *do* because of a common factitive meaning, which is underlying in the operator. After analysing the sequence *which PRP do*, it is evident that there is a partial resemblance to the grammatical *do* present in PAE combined with a partial resemblance to the lexical *do* present notably in *do so* substitutions. It could be postulated that, when the relative subordinate clause is of a positive polarity, *do* is sometimes auxiliary (in particular when it refers to a predicate denoting a purely stative event), sometimes lexical (when it is modified by an adjunct). However, the possibility of stative anaphora with one or several adjuncts indicates an overlap between the grammatical and the lexical, as properties which are typical of lexical *do* are found in grammatical *do*. This is reinforced in utterances where a demonstrative pronoun, substituting for an infinitival predicate, is preposed to an *operator* when *do* is finite (*that* [p\*] [do] (*n't/not*)) and is the object of a *lexical verb* (cf. *do that*) when *do* is non-finite. The study of this second structure confirms the presence of fuzzy category boundaries between grammatical and lexical uses of *do*, which crystallizes in the lexical status of the verb in *this PRP* (finite/non-finite) *do* substitutions. These findings pave the way for a broader study of the verb, expanding the scope of the present research to more complex subject noun phrases (e.g. *which some precedents do* (BNC J77)), to sequences allowing for intervening items such as adverbs (e.g. *which NP apparently do*) and to other cases of

syntactic-semantic ambiguity with *do* (e.g. *I tidied the bedrooms today. When I did yours I couldn't help looking at the photographs in the leather frame.* (BNC JYC)).

## References

- Abbott, Edwin Abbott. 1875. *A Shakespearian Grammar*. London.
- Callaway, Morgan Jr. 1913. *The Infinitive in Anglo-Saxon*. Washington, D.C.: Carnegie Institution.
- Culicover, Peter– Jackendoff, Ray. 2005. *Simpler Syntax*. Oxford: O.U.P.
- Davies, Mark. 2004–. BYU-BNC: The British National Corpus. <http://corpus.byu.edu/bncl>.
- Davies, Mark. 2008–. BYU-COCA: The Corpus of Contemporary American English. <http://corpus.byu.edu/cocal>.
- Denison, David. 1993. *English Historical Syntax*. Harlow: Longman.
- Ellegård, Alvar. 1953. *The Auxiliary 'Do': the Establishment and Regulation of its Growth in English*. Stockholm: Almqvist and Wiksell. (Gothenburg Studies in English, 11)
- Guimier, Claude. 1981. Sur la substitution verbale en anglais. *Modèles Linguistiques*, 3(1), 135–161.
- Hopper, Paul J.–Traugott, Elizabeth. 1993. *Grammaticalization*. Cambridge: C.U.P.
- Huddleston, Rodney D.–Pullum, Geoffrey K. 2002. *The Cambridge Grammar of the English Language*. Cambridge: C.U.P.
- Kroch, Anthony–Pintzuk, Susan–Myhill, John. 1982. « Understanding DO ». In K. Tuite et al. (eds), *Papers from the 18th Regional Meeting of the Chicago Linguistics Society*. Chicago: Chicago Linguistics Society, 282–294.
- Lakoff, George– Ross, James R. 1976. Why you can't do so into the sink. In James D. McCawley (ed), *Notes from the Linguistic Underground*. New York: Academic Press, 101–111. (Syntax and Semantics, 7)
- Méry, Renaud. 2002. À propos du verbe DO. *Anglophonia*, 12, 81–113.
- Miller, Philip. 2011. The choice between verbal anaphors in discourse. *Lecture Notes in Computer Science*, 7099, 82-95. <https://docs.google.com/viewer? a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbncxwaGlsaXBtaWxsZXJsaW5ndWlzdGljc3xneDoyODQwYmI3NDgxNGUxOTAy> (Last accessed 7 March 2019), 1–14.
- Oxford English Dictionary (O.E.D.)*, 2nd edition, 20 vol. 1989. Oxford: O.U.P.

- Quirk, Randolph–Greenbaum, Sidney–Leech, Geoffrey–Svartvik, Jan. 1985. *A Comprehensive Grammar of the English Language*. New York: Longman.
- Royster, James F. 1914–1915. The *Do* Auxiliary – 1400–1450. *Modern Philology*, 12(7), 449–456.
- Royster, James F. 1918. The causative use of *hâtan*. *Journal of English and Germanic Philosophy*, 17, 82–93.
- Royster, James F. 1922. Old English causative verbs. *Studies in Philology*, 19, 328–356
- Sharifzadeh, Saghie. 2012. *Recherche sur Do, Lexique et Grammaire*, PhD dissertation, English studies: Paris-Sorbonne.
- Sharifzadeh, Saghie. 2013. *Do* auxiliaire ou *do* lexical? Indifférenciation sémantique et/ou syntaxique en contexte anaphorique. In *Anglophonia*, 34, 67–87.
- Sharifzadeh, Saghie. 2019. L’anaphore prédicative en anglais: étude comparative des reprises en *do so*, *do it*, *do this*, *do that* et *do* opérateur. *Marqueurs et Structures*, Limoges: Lambert-Lucas.
- Visser, Fredericus T. 1963–1973. *An Historical Syntax of the English Language*, 3 vol. Leiden: E. J. Brill.
- Zilling, Otto. 1918. *Das Hilfsverb “Do” im Mittel-Englischen*. Halle: E. Karras.