A Constraint-Based, Lexicalist Approach to the Periphrastic *Do* Constructions

Jong-Bok Kim
Kyung Hee University
jongbok@nms.kyunghee.ac.kr

November 23, 1999

Abstract

One of the unique properties of English is that it requires the so-called dummy or periphrastic *do* in finite sentential negation constructions. Other negative adverbs such as *never* do not exhibit this requirement. A standard analysis for capturing such a peculiar property has been the so-called *do*-support analysis ever since Chomsky (1957). In such a derivational analysis, the dummy *do* has been added to save an illformed derivation. This paper sketches an alternative, non-derivational account of its properties without resorting to such an insertion mechanism. The constraint-based analysis presented here shows that the enriched lexical representations of *do* and the interaction of the elementary morphosyntactic and valence properties of lexical heads are suffice enough to capture various properties of *do* as well as related phenomena.

1 Basic Properties of *Do*

1.1 Similarities

The supposition of *do* as an extra element that can save a crash in derivation is just like endowing it onto a special status. However, there is ample evidence that this verb is just like the other auxiliaries, at least in terms of syntax. One clear property of auxiliaries is that they have the so called NICE properties, as noted in the literature (cf. Gazdar et al. 1982, Quirk et al. 1985, Warner 1993, inter alia). That is, unlike main verbs, they are sensitive to Negation, Inversion, Contraction and Ellipsis phenomena.

1. Negation: Only auxiliary verbs can be followed by *not* as a sentential negation.

(1) a. Tom will not leave.
b. *Tom kicked not a ball.

2. Inversion: Only auxiliary verbs can undergo the subject-aux inversion.

(2) a. Will Tom leave the party now?
b. *Left Tom the party already?

3. Contraction: only auxiliary verbs can have contracted forms with the suffix *n't*.

(3) a. John couldn't leave the party.
b. *John leftn't the party early.

4. Ellipsis: The complement of an auxiliary verb, but not of a main verb can be elided.

(4) a. If anybody is spoiling the children, John is __ .
b. *If anybody keeps spoiling the children, John keeps __ .

It is not difficult to observe that the auxiliary *do* is also sensitive to these NICE properties.

(5) a. John does not drink alcohol.
b. Does John drink alcohol?
c. John doesn't drink alcohol.
d. John did not take a nap, but Chris did __ .

Various inversion constructions further show that the periphrastic *do* acts just like other auxiliaries. They all are akin to constructions such as emphatic inversions and tag questions:

(6) a. In no other circumstances can John drink alcohol.
b. John could drink alcohol, couldn't he?
c. I could see what was intended, and so could Harry.

(7) a. In no other circumstances does John drink alcohol.
b. John drinks alcohol, doesn't he?
c. I saw what was intended, and so did Harry.

Another similarity comes from the fact that like modal auxiliaries, *do* does not appear in infinitival clauses.

(8) a. *They expected us to do leave him.
b. *They expected us to can leave him.

We can also observe that *do* can be used 'emphatically' like other auxiliaries, in focusing the affirmation or negation of the sentence in question (stress is indicated in the notation by capitalization).

(9) a. He DOES drink alcohol.
    b. He CAN drink alcohol.

These similar properties with auxiliaries give us enough reason to assume that the periphrastic *do* belongs to the same categorial group as auxiliaries.[1] This grouping will predict that *do* appears in the same range of environments as auxiliaries.

1.2 Differences

Despite these similarities, there are some properties that distinguish *do* from other auxiliaries. First, unlike other auxiliaries, *do* appears neither before nor after any other auxiliary:

(10) a. He may be leaving.
     b. He may have been eating
     d. They will have come.

(11) a. *He does be leaving.
     b. *He does have been eating.
     c. *They will do come.

Second, the verb *do* has no obvious intrinsic meaning to speak of. Except for the grammatical information such as tense and agreement, it does not carry any semantic value. Third, if *do* itself is positive, then *do* needs to be emphatic (stressed). But in negative sentences, no such requirement exists.[2]

(12) a. *John does leave.
     b. John DOES leave.

(13) a. John did not come.
     b. John DID not come.

There seems not to be an issue of how we capture the similar properties of *do* and other auxiliaries. For whatever apparatus we adopt for auxiliaries and modals, we can adopt the same one for *do* also. At stake is how we capture the differences from other auxiliaries. The

[1]See Ouhalla (1990) for a treatment of auxiliaries and modals as a category of the functional head Aspect.

[2]But, in what follows we will see that the present analysis predicts the occurrences of *not* in (13)a and (13)b to be different.

differences we have noticed imply that the dummy *do* occurs in more restricted environments than other auxiliaries. Do these properties then require *do* to be introduced by a different mechanism like the English particular rule *do*-support, unlike other auxiliaries? I assume they do not. Instead of adopting this common syntactic rule, I exploit an analysis in which *do* is base-generated. The analysis will argue that its peculiar (distributional) properties, distinct from other auxiliaries, are a reflection of its lexical properties and the interaction among relevant constraints.

2 A Lexicalist, Constraint-Based Approach

2.1 Lexical Entry and Predictions

A simplest way to capture the basic properties of *do* is to assume that the periphrastic *do* has the following lexical entry represented in a format of HPSG's feature structures.

(14)
$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{AUX} & + \\ \text{VFORM} & \text{fin} \end{bmatrix} \\ \text{SUBJ} & \left\langle \boxed{1}\text{NP} \right\rangle \\ \text{COMPS} & \left\langle \boxed{2}\text{VP}\begin{bmatrix} \text{VFROM bse} \\ \text{AUX } - \\ \text{SUBJ } \langle \boxed{1} \rangle \end{bmatrix} \right\rangle \end{bmatrix}$$

The lexical entry in (15) tells that *do* is an auxiliary verb whose verb inflection form (VFORM) is finite, and that it selects a subject NP and a VP whose unsaturated subject is structure-sharing with the subject. Let us consider what this lexical information means in grammar.

[**HEAD** [**AUX** +]]: The lexical entry specifies that like other auxiliaries including modals, *do* is specified to be a verb with the feature [+AUX]. The feature specification [+AUX] ensures that like other auxiliary elements, *do* is also sensitive to negation, inversion, contraction, and ellipsis, as we have observed earlier. This morphosyntactic specification is a simple generalization for NICE properties.

[**HEAD** [**VFORM fin**]]: Unlike auxiliaries *have* and *be* but like modals, *do* is specified to be *fin(ite)*. This property then accounts for why no auxiliary element can precede *do*.

(15) a. He might $_{[bse]}$[have left].
b. *He might $_{[fin]}$[do leave].

Modals like *might* select a base VP. But in (15)b *might* combines with a finite VP headed by the finite *do*. This feature specification further explains why *do* cannot appear in infinitival clauses:

(16) a. John believed Kim to $_{[bse]}$[have left here].
b. *John believed Kim to $_{[fin]}$[do leave here].

COMPS <VP[bse, −AUX]>: As its valence information, *do* requires a subject NP and a VP. The requirement on the complement VP is [*bse*]. This feature specification blocks modals from heading the VP following *do*. Since modals are specified to be [*fin*], the ungrammaticality of (17) is a natural expectation.

(17) a. *He do $_{[fin]}$[can leave here].
b. *He do $_{[fin]}$[may leave here].

This restriction also accounts for the inflection of the verb following *do*:

(18) Pat did not [go/*went/*goes/*gone/*going to the conference].

Its complement VP is also required to be [−AUX]. This specification will correctly predict the ungrammaticality of examples in (19) and (20).

(19) a. *Jim [DOES $_{[+AUX]}$[have supported the theory]].
b. *The proposal [DID $_{[+AUX]}$[be endorsed by Clinton]].

(20) a. *I [do [not $_{[+AUX]}$[have sung]]].
b. *I [do [not $_{[+AUX]}$[be happy]]].

In (19) and (20), the VPs following the auxiliary *do*, stressed or not, bear the feature [+AUX] inherited from the auxiliaries *have* an *be*. This explains their ungrammaticality.

Raising Properties: Further, like other auxiliaries, *do* selects a subject NP and a VP complement whose unrealized subject is structure-shared with its subject ($\boxed{1}$). Treating *do* as a raising verb like other English auxiliaries is based on typical properties of raising verbs that differentiate them from equi verbs. (cf. Pollard & Sag 1994): (a) raising verbs, unlike equi verbs, do not by themselves assign any semantic role to their subject. (b) the index of the role-assigned subject in equi verbs should be 'referential', but no such restriction appears on the subject of raising verbs, and (c) unlike equi verbs, raising verbs do not allow NP complements. Auxiliaries including *do* have these raising verb properties (See Pollard and Sag 1994).

(21) a. John may leave.
b. It may rain.
c. *John/*It may something.

(22) a. John did not leave.
b. It did not rain.
c. *John/*It did not something.

No Intrinsic Meaning: Though *do* historically may be derived from the causative *do*, it has no intrinsic semantics in modern English (see Pollock 1989). The difference in the CONT(ENT) value of auxiliaries like *can* and that of *do* given in (23) illustrates this point.

(23) a. can:

$$\left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{l} \textit{verb} \\ \text{AUX} + \\ \text{VFORM fin}\end{array}\right] \\ \text{SUBJ} & \langle \boxed{1}\text{NP}\rangle \\ \text{COMPS} & \left\langle \text{VP}\left[\begin{array}{ll}\text{HEAD} & [\text{VFORM bse}] \\ \text{SUBJ} & \langle\boxed{1}\rangle \\ \text{CONT} & \boxed{2}\end{array}\right]\right\rangle \\ \text{CONT} & \left[\begin{array}{ll}\text{RELN} & \textit{can} \\ \text{SOA-ARG} & \boxed{2}\end{array}\right]\end{array}\right]$$

b. do:

$$\left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{l} \textit{verb} \\ \text{AUX} + \\ \text{VFORM fin}\end{array}\right] \\ \text{SUBJ} & \langle \boxed{1}\rangle \\ \text{COMPS} & \left\langle \text{VP}\left[\begin{array}{ll}\text{HEAD} & \left[\begin{array}{l}\text{VFORM bse} \\ \text{AUX} -\end{array}\right] \\ \text{SUBJ} & \langle\boxed{1}\rangle \\ \text{CONT} & \boxed{2}\end{array}\right]\right\rangle \\ \text{CONT} & \boxed{2}\end{array}\right]$$

Although like other auxiliaries *do* is treated as a raising verb, it is lexically specified to have no semantic relation: the structure sharing of its CONTENT value and its VP complement's CONTENT value ($\boxed{2}$) guarantees this.[3]

The lexical entries in (23) show the similarities and differences between *do* and modals including *can*. They all are [+AUX] and finite in terms of head features. They are also alike in that they select a VP[bse] whose unsaturated subject is structure-sharing with its own subject (i.e., raising properties). The difference lies in their semantic content and a further specification on the VP complement of *do*: it should be [−AUX]. The enriched lexical information of *do* captures the basic properties of *do* in a straightforward manner.

[3] The lexical entry of *do* is thus similar to that of *to*, in that they both are treated as raising verbs and their meanings are identical to those of their VP complements. Pullum (1982) notes that *to* and *do*, in addition to differing by one phonological feature, *voicing*, differ in one small way: *do* appears only in finite contexts, and *to* only in non-finite contexts. Other than that, they share the property that they obligatorily take bare verbal complements (hence not modals), which only have finite forms.

2.2 Stressed vs. Unstressed *do*

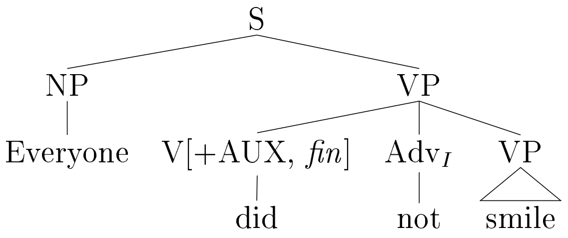
One remaining property of the dummy *do* we haven't discussed is that if *do* itself is positive, then *do* needs to be emphatic (stressed):

(24) a. *John did answer the question.
b. John DID answer the question.
c. John did not answer the question.

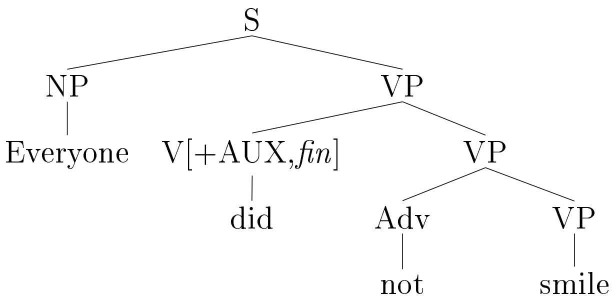
Though there is a question of whether or not the grammar needs to block the unstressed *do* in positive declaratives, I here sketch a lexical account.

The gist of the proposed analysis starts from introducing the analysis of Kim and Sag (1995) and Kim (2000) in which the negative marker *not* lives a double life. One is an adverb that negates the nonfinite VP constituent that it appears to the left of, while the other is introduced by a lexical rule as a complement of a finite auxiliary verb and negates the whole clause. This could be the source of the ambiguity of sentences like *Everyone did not smile.* as represented in (25).

(25)



(26)



In (25), *not* is a complement of *did* and has scope over the entire clause, while in (26), it is adjoined to the VP *smile* and has scope only over the lower VP.[4]

The constraint we want to add is a phonological condition. In particular, I impose a phonological restriction on the output of the Conversion Lexical Rule as given in (27).

[4]Kim and Sag (1995) and Kim (2000) show that this lexical analysis can better account for the properties of English negation and related phenomean such as VP ellpsis and for the systematic differences between English and French.

(27) English (Negation) Conversion Lexical Rule:

$$\mathrm{V}[+\mathrm{AUX},\ \textit{fin}] \begin{bmatrix} \mathrm{COMPS} & \mathrm{L} \\ \mathrm{CONT} & \boxed{2} \end{bmatrix} \Rightarrow \mathrm{V}[+\mathrm{AUX},\ \textit{fin, unstressed}] \begin{bmatrix} \mathrm{COMPS} & \left\langle \mathrm{ADV}_I\text{: } \boxed{3} \right\rangle \oplus \mathrm{L} \\ \mathrm{CONT} & \boxed{3}\begin{bmatrix} \mathrm{ARG} & \boxed{2} \end{bmatrix} \end{bmatrix}$$

The main effect of the output of this lexical rule is to allow a structure like the one in (25). This lexical rule takes as input any finite auxiliary which selects for a base VP complement and yields as output another verbal entry which adds an Adv_I adverb (such as *not*) as an additional complement, i.e., adds it onto the finite verb's COMPS list.[5]

The main motivation of adding the condition ([*unstressed*]) on the output of the lexical rule concerns the scope relation between the head verb and its added complement *not*. As noticed, the output semantic content of the lexical rule specifies that the added complement *not* takes wide scope over the head. But notice that this semantic condition holds only when the auxiliary verb is not stressed.[6]

(28) a. He CAN not go to school tomorrow, (can't he/*can he)?
b. He MAY not go to school tomorrow, (mayn't he/*may he)?

The possible type of tag questions in (28) shows that the negator *not* following the stressed auxiliaries does not have wide scope. The dummy *do* is not different in this respect. Examples in (29) illustrate this:

(29) a. He DID not go to school yesterday, (didn't he/*did he)?
b. He DID not come, and so did she/*neither did she.

The introduction of the condition predicts that the negator following the unstressed *do* always takes a wide scope. The test of a tag construction again can prove this:

(30) a. He cannot attend and neither can she.
b. *He cannot attend and so can she.

(31) a. He did not come and neither did she.
b. He did not come and *so did she.

[5] Adv_I restricts adverbial complements to only a small subset of adverbs like *not* and possibly *so* in English. The lexical rule also has a semantic effect: the converted complement adverb including negation takes the meaning of the input verb as its argument, as can be seen from the output CONTENT value. See Kim and Sag (1995) and Kim (2000).

[6] Such a restriction on auxiliaries and modals seems to hold only in declarative negative sentences, not in questions or ellipsis. Stress on an auxiliary in these constructions appears not to affect its scope.

Given that the narrow scope negation triggers the *so* tag, whereas the wide scope triggers the *neither* tag, the unacceptability of the *so* tag in (30)b and (31)b shows that *not* here takes wide scope. This scope fact is a direct consequence of the lexical rule application.

But *do* is different from other auxiliaries in one important respect, as noted earlier: it should be stressed if not followed by a sentential negation among other things. I assume that this requirement is due to a blocking effect. Blocking is a phenomenon whereby the availability of a better-suited or more specific form renders a less specific one ungrammatical. Consider the examples in (32)[7].

(32) a. *He did walk.
b. He walked.

The semantically empty verb *did* in (32)a, if not assigned stress for its emphatic usage, has no function at all other than the realization of tense information. To account for why the existence of *walked* blocks the phrase *did walk*, one can resort either to a pragmatic effect, or to a morphological blocking effect. In terms of a pragmatic approach, *did come* would be blocked, since English speakers choose the simpler form of expressing the same function.[8] Instead of relying on the notion of minimization of effort or least effort, one can also have a theory of a morphological blocking with the extension of the domain of blocking to a phrasal level, as proposed by Poser (1992). In this spirit, the lexical instantiation of tense information would block its instantiation at a phrasal level. More specifically we can assume the following condition:[9].

(33) Tense Realization in English:
In English, tense is realized at a lexical level (by a morphological element) rather than at a phrasal level (by an independent word).

The occurrence of *do* in (32)a, if unstressed, has only the function of instantiating the tense information.[10] Since there exists the lexical form *walked* where the tense is realized as an affix, the periphrastic form *did* is blocked.[11]

[7]In the accounts of Lapointe 1980, Gazdar et al. 1982, Falk 1984, and Pollock 1989, among others, the unstressed *do* in declarative sentences as in *they do go* is not explicitly blocked.

[8]See Poser 1992 for a brief note on a pragmatic approach.

[9]A similar condition is, implicitly or explicitly, also assumed in Hudson 1976, Gazdar et al. 1982, Warner 1993, among others

[10]One can wonder if the general condition is that the existence of morphologically simple forms blocks the well-formedness of syntactically complex expressions with the same semantic content, why then are contractions permitted at all. Given the *n't* is an inflectional marker, *Tom did not leave* should be blocked by the existence of *Robin didn't leave*. We could attribute this coexistence to a pragmatic difference: formal and informal style.

[11]The same method can be applied to English comparative adjectives: the morphological category of comparative adjective can be either a lexical form or a periphrastic form. Thus if there exists a lexical form, the periphrastic form is blocked. For further details, see Poser 1992.

The sketched analysis thus predicts the unacceptability of (34) in which the unstressed *do* is followed by elements other than *not*.

(34) a. *Kim does [never [eat bagels]].
b. *Kim does [probably [leave]].

The condition on tense realization in English prevents the unstressed *do* in (34). If unstressed, *do* in (34)a is blocked by the existence of *Kim never eats bagels.* in which the tense is realized on the verb. But the stressed *DO* in (35) has no such requirement.

(35) a. Kim DOES never eat bagels.
b. Kim DOES probably leave.

The two key points of the assumed analysis are thus the phonological condition ([*unstressed*]) on the output of the lexical rule and the 'stress' requirement on *do* in positive declarative sentences by a blocking effect. These two mechanisms ensure that the unstressed *do* in declarative cases takes *not* as a complement, and that *do* in declarative sentences is stressed. Consider the contrast in (36) and (37) again.

(36) a. *John did come
b. John DID come.

(37) a. John [did] [not] [come].
b. John DID [not come], (didn't he?).

The contrast between (36)a and (36)b is due to the blocking effect given in (32). The unstressed *do* in (37)a, the output of the lexical rule, selects *not* as its complement. Also, since the unstressed *do* selects *not* as its complement, *do* in (36)a does not satisfy its subcategorization requirement. The stressed *DID* is also possible preceding the negator *not*, as in (37)b. But notice that *not* here cannot be a complement. It can be only a modifier.

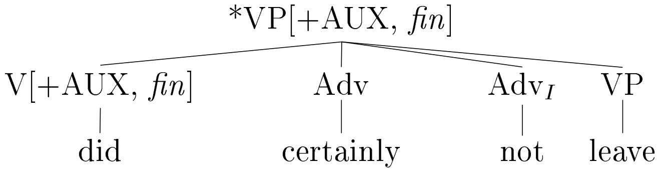
The analysis suggested here can easily explain the contrast between (38) and (39).

(38) a. John certainly can not leave.
b. John can certainly not leave.

(39) a. John certainly did not leave.
b. *John did certainly not leave.

In the proposed analysis, *not* in (38)a can be either a complement or a modifier, whereas *not* in (38)b can only be a modifier. But the revised lexical rule ensures that the unstressed *did* in (39) selects *not* as its complement, allowing examples like (39)a. But (39)b is ruled out because the VP modifier *certainly* intervenes between AUX and the complement *not*, as shown in (40).[12]

(40)

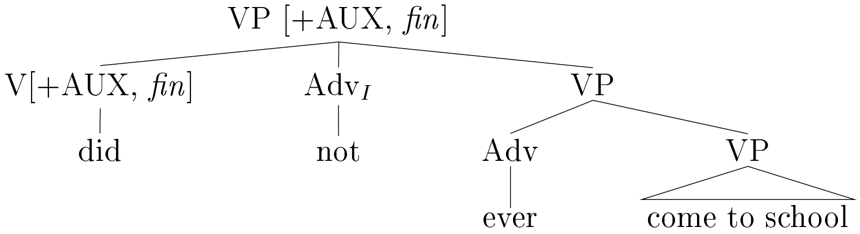


The proposed analysis, however, does not prevent us from generating the example like (41).

(41) John did not ever come.

The unstressed *did* select the negative marker *not* as its complement in addition to a VP. And the VP modifier *ever* modify this VP complement, as represented in (42).

(42)



Our analysis thus can provide a simple answer to the contrast in (39) and (40).

In sum, the unstressed *do* in each case possesses at least one additional property that makes it not a dummy, but an independent word with a certain function other than tense realization:

(43) a. He DID walk.
b. He did not walk.
c. Did he walk?
d. Mary didn't walk, but Tom did __ .

[12] An analysis such as an English particular rule that moves unstressed finite verbs to the left periphery of their phrases, as adopted in Baker (1991), will also encounter a problem in accounting for such a contrast (cf. Baker 1991, fn 14). The rule should be able to move the unstressed *did* in (39)a to the left of the adverb *probably*.

(i) a. Nora probably did not ever open the letter.
b. *Nora did probably not ever open the letter.

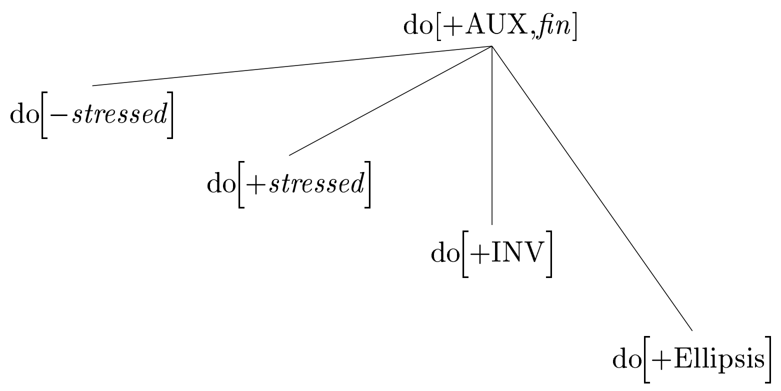
When *do* is stressed as in (43)a, it serves as a word for a contrast of polarity. In (43)b, it is a verb selecting *not* as a complement. In (43)c, it has one additional feature [+INV] that allows the question. In (43)d, *do* is the verb that triggers the ellipsis, and this *do* is further the only available element on which tense can be realized.[13] Thus in each occurrence here, *do* serves as an independent word with its own specific function. It is not just a sort of instantiation of tense information.[14] The difference in the valence information of each usage of *do* is represented more formally in (44):

(44) a. $\text{do}\begin{bmatrix} -\textit{stressed} \\ \text{COMPS} < \text{Adv}_I, \text{ VP[bse]} > \end{bmatrix}$ b. $\text{do}\begin{bmatrix} +\textit{stressed} \\ \text{COMPS} < \text{VP[bse]} > \end{bmatrix}$

c. $\text{do}\begin{bmatrix} +\text{INV} \\ \text{COMPS} < \text{NP[nom],VP} > \end{bmatrix}$ d. $\text{do}\begin{bmatrix} \text{SUBJ} < \boxed{1}\text{NP} > \\ \text{COMPS} < > \\ \text{ARG-ST} < \boxed{1}, \text{ VP} > \end{bmatrix}$

This does not mean that each *do* has no connections at all. As represented in the hierarchy in (45), they all share the information on the auxiliarihood and finiteness.[15]

(45)



2.3 *Do* in Imperatives

Note that there are differences between *do* in imperatives and *do* in non-imperatives. One telling difference is that *do* in imperatives can occur before another auxiliary like *be* and *have*.

[13] For an analysis of VP Ellipsis within the framework of HPSG, see Kim (2000).

[14] Thus there is no stress condition on the dummy *do* in (43)c and (43)d. In the present analysis, those exceptional cases or middle English where *do* occurs without the stress in declarative clauses do not observe the Tense Realization Condition in English.

[15] For the functions of multiple inheritance hierarchy, see Pollard and Sag (1994) and Sag and Wasow (1999).

(46) a. Do be honest!
b. Do have reached a decision regarding the matter!

Do in imperatives places no restriction on the auxiliary value of its VP complement. This lose restriction can be incorporated into our analysis with a minor modification to the lexical entry for *do* in nonimperatives:

(47)
$$\begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{AUX} & + \\ \text{MOOD} & \text{imperative} \end{bmatrix} \\ \text{SUBJ} & \langle \quad \rangle \\ \text{COMPS} & \left\langle \boxed{1}\text{VP[SUBJ } \langle \boxed{2}\text{NP[PER 2nd]} \rangle] \right\rangle \\ \text{ARG-ST} & \left\langle \boxed{2}, \boxed{1} \right\rangle \end{bmatrix}$$

The lexical entry in (47) specifies that *do* in imperatives selects a VP whose subject is 2nd in person. This subject is syntactically empty but represented in the argument structure. The value of ARG-ST (argument-structure) includes the subject and the VP complement of *do*. This will guarantee the correct semantics as well as account for binding relations (defined within the domain of ARG-ST) in imperative sentences like *Wash yourself/*himself.*[16] Given this minimal difference in lexical information between *do* in imperatives and the one in nonimperatives, it is not difficult to predict the contrast such as the one in (48).

(48) a. *They DO $_{[+AUX]}$[have been eating].
b. Do $_{[+AUX]}$[have some more tea]!

Though *do* in imperatives cannot have subject as in **Do you sit down!*, *don't* allows the subject as in (49). One interesting property in forming a negative imperative is that unlike *don't* the two-word sequence *do not* cannot be used with imperatives containing an overt subject (See Postman 1996 for a detailed discussion).

(49) a. Don't you sit down over there!
b. Don't anybody say anything!

(50) a. *Do not you sit down over there!
b. *Do not anybody say anything!

[16]The binding principle of HPSG (Sag and Wasow (1999) states that a reflexive pronoun must be bound by a preceding argument of the same verb.

This contrast tells that *do* and *don't* have different lexical information from those in nonimperatives. This could be easily accounted for within an analysis where *n't* is taken to be an inflection (cf. Zwicky and Pullum 1983). A strong argument to treat *n't* as an nonproductive inflectional element rather than as an productive syntactic element can be seen from its lexical idiosyncrasies. Not all combinations of auxiliaries with *n't* are acceptable as in **willn't, *amn't*, and **mayn't*. Accepting this reasoning, we assume that *don't* and *do* in *do not* are specified with different lexical information. For example, *don't* in imperatives may have a lexical entry like the following:

(51)
$$\begin{bmatrix} \text{I-FORM} & \text{don't} \\ \text{HEAD} & \begin{bmatrix} \textit{verb} & \\ \text{INV} & + \\ \text{AUX} & + \\ \text{MOOD} & \text{imperative} \\ \text{NEG} & + \end{bmatrix} \\ \text{SUBJ} & \langle\quad\rangle \\ \text{COMPS} & \left\langle (\boxed{2}), \text{VP[SUBJ} \left\langle \boxed{2}\text{NP} \right\rangle] \right\rangle \end{bmatrix}$$

The lexical entry specifies that *don't* in imperatives is always inverted and can optionally combine with its unsaturated VP's subject.

Our analysis requires no additional mechanism to account for the ungrammaticality of cases like (50). Unlike *don't* in imperatives, *do* cannot combine with a subject followed by a VP as its syntactic sisters. It only takes a VP complement. The contrast in (52) could also be a direct prediction of the analysis presented here.

(52) a. *Do you not desert me!
b. Do you not like artichokes?

Do in imperatives cannot select for an overt subject as its lexical information; however nothing is wrong for an inverted *do* in an inverted environment to select a subject and a base VP which is in turn modified by the constituent negation *not* in (52)b.

Though imperative *do*, *don't*, and *do not* are different from those in tensed nonimperatives, they share the properties of auxiliarihood (Data (53)–(54) from Postman 1996: p178), in that they all bear the feature value [+AUX]. This common property explains the possibility of VP ellipsis:

(53) a. Boys will taste the food if girls do ___ .
b. Boys like the food but girls do not ___ .
c. Boys like the food but girls don't ___ .

(54) a. Did we say you could draw on the walls? All right, then, don't __ !
b. We want everyone to come, so those who can, by all means do __ !
c. [I'm going to open the oven and peek at the pie.] Do not __ ! You'll ruin it.

Given the VP ellipsis is licensed by an auxiliary element (cf. Kim and Sag (1995) and Kim (2000)), the well-formedness of these ellipsis sentences is a natural consequence.

The auxiliary *do* is realized in various environments with different lexical realizations. In imperatives, we observed two instances: *don't* and *do*, in addition to at least the four usages of *do* in nonimperatives. What we have observed so far is that each usage has its own function and unique properties. However, this does not mean that there is no sharing properties. As observed, one topmost property that all instances of *do* bear is the auxiliarihood, as represented in the following hierarchy.

(55) do[+AUX]

$$\text{do}\begin{bmatrix}\text{VFORM fin}\\ \text{MOOD indicative}\end{bmatrix} \qquad \text{do}\begin{bmatrix}\text{VFORM base}\\ \text{MOOD imperative}\end{bmatrix}$$

$$\text{don't}\big[+\text{NEG}\big] \qquad \text{do}\big[-\text{NEG}\big]$$

Such an analysis is different from an analysis such as that of Pollock (1989). Pollock (1989) accepts the imperative *do* or *don't* is a main verb unlike those in indicative clauses. This move will obviously miss the common properties of *do* in every environment. For example, Every *do* is sensitive to the VP ellipsis phenomenon. No such problem arises in our auxiliary analysis.

3 Brief Comparison with a *Do*-support Approach

As noted earlier, one of the main mechanisms for capturing the peculiar properties of *do* has been the *do*-support rule, originating in Chomsky (1957) and revived in Chomsky (1991). The major criterion for inserting *do* in syntax has been the 'nonadjacency' of an abstract Tense element with the main verb, as represented in (56).

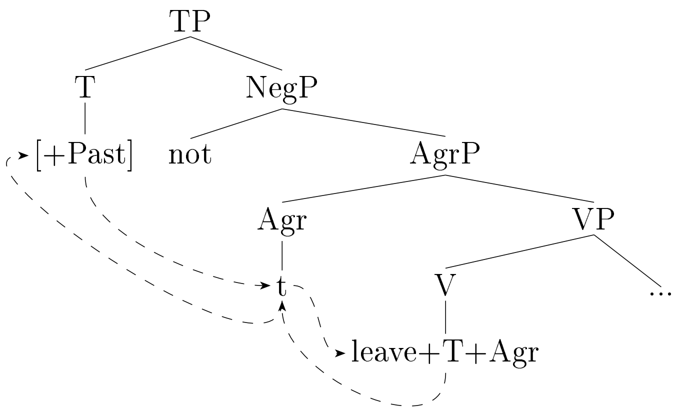
(56) a. Past John swim? → Did John swim?
b. John Past not swim. → John did not swim.
c. John Past Emph swim. → John DID swim.

This insertion analysis relying on the notion of 'adjacency' suffers from some problems. It first meets a difficulty in accounting for cases like (57)a,b where adverbs intervene between the unstressed *do* and *not*.

(57) a. *Nora did probably open the letter.
b. *Nora did never open the letter.
c. Nora did not open the letter.

Though in (57)a,b, the tense and the main verb are separate, no *do*-support takes place. The contrast here may be stated in terms of differentiating *not* from other adverbs, as in Pollock (1989) and Chomsky (1991). In Chomsky's (1991) analysis where *not* is taken to occupy the head of NegP, *do*-insertion is forced by the inability of I to lower to the main verb:

(58)



According to the derivational structure in (58), Tense is lowered to Agr and then its complex is lowered to V at s-structure in order to get the inflected main verb *left*, as in *John left.* Since this process leaves an ungoverned trace of Tense, Chomsky's analysis reraises the complex V-Tense-Agr to Tense again at LF. In this reraising process, the trace of Agr is deleted under the assumption that unnecessary elements are deletable. This deletion will then escape the ECP violation of LF reraising. But when the Neg head is filled with *not*, this reraising is impossible, since the head Neg causes the violation of the HMC (Head Movement Constraint) and hence ECP. Chomsky claims that the last resort that can save such a derivation is *do*-insertion, which can eventually generate sentences like *John did not leave*, but blocking sentences like **John not left.*

Such an LF movement analysis correctly predicts the contrast between (59)a and (59)b.

(59) a. *John not often cleaned the room.
b. John did not often clean the room.

Not in (59)a is the key factor for the violation of the HMC whereas *did* in (59)b can save such a derivation. But consider examples like (60).

(60) a. John will often not attend the meetings.
b. John will probably not open the letter.

The most likely position of the adverbs in (60) will be NegP-adjoined, as represented in (58). One immediate question, then, arises as to why *do*-support does not render sentences like (61)a grammatical (see Battistella (1987) for a similar point).

(61) a *John did often not attend the meetings. (unstressed *do*)
b.*John did probably not open the letter. (unstressed *do*)

If the language particular rule *do*-support plays a crucial role in saving a crashed derivation, there seems to be no obvious reason why it cannot happen for cases like (61)a and b. But the present analysis, as we have seen, can provide a straightforward way of capturing this contrast: the VP adverbs intervene between the lexical head, unstressed *do* and its complement *not*. This explains the ungrammaticality of (61).

4 Conclusion

The dummy *do* has dual properties; auxiliary-like properties and non-auxiliary like properties. We have seen that with respect to the NICE properties, the verb behaves just like other auxiliaries. But its distributional behavior and semantics places it into a different category.

This paper has provided a simple, lexicalist treatment capturing these dual properties. There are three main points of this analysis. The first is the lexical entry for *do*. We have assumed the verb is a finite auxiliary verb selecting for an non-auxiliary verb phrase as its complement. This lexical specification allowed us to capture its various distributional restrictions as well as basic properties. Another gist of this analysis is that following Kim and Sag (1995) and Kim (2000) , it allows the unstressed *do* verb to take the negative marker *not* to be its complement. This analysis, motivated from other independent phenomena such as VP ellipsis (see Kim 2000 for detail) could predict the phonological behavior of *do*. The final main point of our analysis is the Tense Realization condition in English. This blocking condition accounts for the requirement of stressing *do* in a positive while blocking an unstressed *do* from occurring in a positive sentence.

We have observed that the auxiliary *do* bears numerous idiosyncratic properties in addition to the properties of true auxiliary verbs like modals. Once we enriched the lexical information of *do* with independently motivated constraints, the interaction of the elementary morphosyntactic and valence properties of lexical heads has become the main source for the clean and streamlined analysis of *do* and related phenomena.

References

Akmajian, Adrian, Susan Steele, and Thomas Wasow. 1979. The Category AUX in Universal Grammar. *Linguistic Inquiry* 10: 1–64.

Baker, C. L. 1991. The Syntax of English *Not*: The Limits of Core Grammar. *Linguistic Inquiry* 22: 387–429.

Chomsky, Noam. 1957. *Syntactic Structures.* The Hague: Mouton.

Chomsky, Noam. 1991. Some Notes on Economy of Derivation and Representation. In R. Friedin (ed.), *Principles and Parameters in Comparative Grammar,* 417–454. Cambridge, MA: MIT Press.

Ernst, Thomas. 1992. The Phrase Structure of English Negation. *The Linguistic Review*, 9: 109–144.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, and Ivan Sag. 1985. *Generalized Phrase Structure Grammar.* Cambridge, MA: Harvard University Press.

Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar.* Cambridge, MA: MIT Press.

Kim, Jong-Bok. 1995. English Negation from a Non-Derivational Perspective. In Jocelyn Ahlers, Leela Bilmes, Joshua S. Guenter, Barbara A. Kaiser, and Ju Namkyung (eds.), *Proceedings of the 21st Annual Meeting of the Berkeley Linguistics Society (BLS)*, 186–197. Berkeley, CA: Berkeley Linguistic Society.

Kim, Jong-Bok. 2000. *The Grammar of Negation: A Constraint-Based Approach.* Stanford: CSLI Publications.

Kim, Jong-Bok and Ivan A. Sag. 1995. The Parametric Variation of French and English Negation. In the *Proceedings of the 14th West Coast Conference on Formal Linguistics (WCCFL XIV).*

Ouhalla, Jamal. 1990. Sentential Negation, Relativized Minimality and the Aspectual Status of Auxiliaries. *The Linguistic Review* 7: 183–231.

Pollock, Jean-Yves. 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20: 365–424.

Pollard, Carl and Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar.* Chicago: University of Chicago Press.

Poser, Bill. 1992. Blocking of Phrasal Constructions by Lexical Items. In Ivan Sag and Anna Szabolcsi (eds.), *Lexical Matters*, 111–130. Stanford: CSLI Publications.

Potsdam, Eric. 1996. Syntactic Issues in the English Imperative. Doctoral Dissertation. University of California, Santa Cruz.

Pullum, Geoffrey K. and Deirdre Wilson. 1977. Autonomous Syntax and the Analysis of Auxiliaries. *Language* 53: 741–788.

Sag, Ivan and Tom Wasow. 1999. *Syntactic Theory: A Formal Introduction.* Stanford: CSLI Publications.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language.* London: Longman.

Warner, Anthony. 1993. *English Auxiliaries: Structure and History.* Cambridge: Cambridge University Press.

Warner, Anthony. 1998. English Auxiliaries without Lexical Rules. Ms. University of York.

Zwicky, Arnold M. 1985. Heads. *Journal of Linguistics* 21, 1–29.

Zwicky, Arnold M. and Geoffrey K. Pullum. 1983. Cliticization vs. Inflection: English n't. *Language* 59: 502–513.

Jong-Bok Kim
Dept. of English
Kyung Hee University
1 Hoegi-dong, Dongdaemoon-gu, Seoul, Korea
Tel: 961-0892/966-8336