

Syntactic Form and Discourse Accessibility

Gregory Ward Andrew Kehler
Northwestern University* University of California San Diego†

Fourth Discourse Anaphora and Anaphora Resolution Colloquium
Estoril, Portugal
September 19, 2002

1 Introduction

Three types of ‘identifier *so*’ (Bolinger 1972):

- (1) “And with complete premeditation [they] resolved that His Imperial Majesty Haile Selassie should be strangled because he was head of the feudal system.” He was *so strangled* on Aug. 26, 1975, in his bed most cruelly. (*Chicago Tribune* 12/15/94)
- (2) Robert D. Krebs, Santa Fe’s chairman, argues that since its businesses are valued in different ways, “the sum of the parts may be greater than the whole.” But it isn’t clear why that should *be so*. (*Wall Street Journal* 11/1/89)
- (3) As an imperial statute the British North America Act could be amended only by the British Parliament, which *did so* on several occasions. (*Groliers Encyclopedia*)

In this talk, we:

- review the heretofore mysterious anaphoric properties of identifier *so*;
- provide a unified and compositional account of both preverbal *so* and *do so*;
- respond to and counter Fu, Roeper, and Borer’s (2001) recent criticism of our earlier account.

2 Properties of Identifier ‘So’

Identifier *so* is used to establish reference to a contextually salient event of the type denoted by the verb it modifies.

Uses of *so* that are **not** identifier:

- preposed propositional *so*: So it seems. So you say. So it is.
- postverbal propositional *so*: I think/suppose/say/believe *so*.
- veridical *so*: Is that *so*?
- consequential *so*: A: He’s a pig. B: So you’re **not** going out with him after all?

- particle *so*: So, how long have you been at Northwestern?
- ‘queer’ *so*: ‘I’ve come to the conclusion,’ he told me, ‘that I’m not really “so” at all. I much prefer girls.’ At this date the cant word among homosexuals for their proclivities was ‘so’. (Mackenzie 1963, as cited in the OED)
- additive *so*: Bill gave a speech and *so* did Hillary.
- intensifier *so*: There are so many uses of ‘so’!
- generation X *so*: I’m like, *so* going out with him.
- and so on and so forth.”

Preverbal and postverbal *so* (Kehler and Ward 1999):

- (4) Oprah spent two weeks studying law to prepare for her court **cas**
 - a. By *so* studying, she was able to handle herself well on the witness stand.
 - b. She studied *so* in order to handle herself well on the witness stand.

1 Preverbal *so*

- (5) “And with complete premeditation [they] resolved that His Imperial Majesty Haile Selassie should be strangled because he was head of the feudal system.” He was *so strangled* on Aug. 26, 1975, in his bed most cruelly. (=1)
- (6) In fact, in substantiating these fears, Judge Bork again essentially concedes that economic freedom is a component of the Constitution: “We already have clauses that could be used to protect economic freedom – and were *so used*.” (*Wall Street Journal*)
- (7) In fact, it is interesting that, in English, at least, there is virtually no marking of an NP with respect to the Discourse-status of the entity it represents. Of course, if an NP is indefinite and is thereby understood as evoking something Hearer-New, we can infer Discourse-New. However, if it is not *so marked*, then, with one exception, we cannot tell from its form whether it has occurred before in the discourse. (Prince 1992, p. 304)

he referent event of the preverbal *so* construction:

- must be HEARER-OLD in the sense of Prince (1992):
 - (8) His Imperial Majesty Haile Selassie was head of the feudal system.
 - a. #He was *so* strangled on Aug. 26, 1975, in his bed most cruelly.
 - b. He was strangled on Aug. 26, 1975, in his bed most cruelly.
 - (9) In fact, in substantiating these fears, Judge Bork again essentially concedes that economic freedom is a component of the Constitution:
 - a. #“We already have clauses that could be *so* used by opponents of the administration’s trade policy.”
 - b. “We already have clauses that could be used by opponents of the administration’s trade policy.”

*Department of Linguistics, 2016 Sheridan Road, Evanston, IL 60208, gw@northwestern.edu

†Department of Linguistics, 9500 Gilman Drive #0108, La Jolla, CA 92093-0108, kehlerr@ling.ucsd.edu

- must be SALIENT:

- (10) With complete premeditation, they resolved that His Imperial Majesty Haile Selassie should be strangled because he was head of the feudal system. They also resolved to commit a variety of other violent acts, although those would come somewhat later. Obviously, these people were very prone to violence.
- a. #Selassie was so strangled on Aug. 26, 1975, in his bed most cruelly.
 - b. Selassie was strangled on Aug. 26, 1975, in his bed most cruelly.

- must be LINGUISTICALLY EVOKED:

- (11) [A and B together have just witnessed Haile Selassie being murdered by strangulation]
- a. A: #He was so strangled most cruelly.
 - b. A: He was strangled most cruelly.
- (12) [A and B have just witnessed a vote in Congress that repealed an amendment of the Constitution]
- a. A: #By so changing the constitution, Congress is setting a dangerous precedent.
 - b. A: By changing the constitution, Congress is setting a dangerous precedent.

- may be INFERRABLE in the sense of Prince (1981, 1992):

- (13) Regarding a possible Elvis Presley stamp, Postmaster General Frank notes that anyone *so honored* must be “demonstrably dead” for 10 years. (*Wall Street Journal*)

2.2 Do so

Previous accounts of the *do so* construction have noted its seemingly idiosyncratic syntactic and anaphoric properties (Lakoff and Ross 1966, Anderson 1968, Bouton 1970, Halliday and Hasan 1976, Hankamer and Sag 1976, Sag and Hankamer 1984, Quirk et al. 1985, Miller 1990, Ward, Sproat, and McKoon 1991, Cornish 1992, Fu and Roeper 1993, Dechaine 1994, Fu, Roeper, and Borer 2001).

- (14) Sam sold his stock on insider information, and Martha did so too.

2.2.1 Syntactic and semantic properties of *do so*

The *do* of *do so* is main verb *do* and not auxiliary *do* (Hankamer and Sag 1976, Quirk et al. 1985, Miller 1990, Dechaine 1994, inter alia):

- main verb *do* does not undergo auxiliary inversion, unlike auxiliary *do* (Miller 1990):

- (15) a. Hillary did so.
 b. *Did Hillary so?
 c. Did Hillary do so?
 d. Hillary did.
 e. Did Hillary?

- main verb *do* permits *so*, unlike auxiliary *do*:

- (16) a. Dubya filed a lawsuit, and Al did too. [auxiliary *do*]
 b. Dubya has filed a lawsuit, and Al has too.
 c. Dubya will file a lawsuit, and Al will too.

- (17) a. Dubya filed a lawsuit, and Al did so too. [main verb *do*]
 b. *Dubya has filed a lawsuit, and Al has so too.
 c. *Dubya will file a lawsuit, and Al will so too.

- *do so* (and its variant *so doing*) is limited to (non-stative) events, unlike auxiliary *do* (Lakoff and Ross 1966, Anderson 1968, Bouton 1970, Hankamer and Sag 1976, Quirk et al. 1985, Miller 1990, Dechaine 1994, inter alia):

- (18) a. Al hates George, and Tipper does too. [auxiliary *do*]
 b. Al hates George, and so does Tipper. [auxiliary *do*]
 c. ??Al hates George, and Tipper does so too. [main verb *do*]
 d. ??Al hates George, and in so doing, is not popular with Republicans. [main verb *do*]

The *so* of *do so* is an adverb (Bouton 1970, Quirk et al. 1985, Hankamer and Sag 1976):

- it doesn't passivize:

- (19) a. *...and so was done by Hillary.
 b. ...and it was done by Hillary.
 c. ...and that was done by Hillary.

- it doesn't cleft:

- (20) a. It is that which Hillary did.
 b. *It is so which Hillary did.
 c. What Hillary did was that.
 d. *What Hillary did was so.

2.2.2 A compositional analysis of *do so*

In our account, *do so* consists of an intransitive, semantically bleached verb *do* with an adverbial *so* modifier. Although postverbal, the *so* of *do so* has precisely the same semantic/pragmatic properties of (preverbal) identifier *so*, paralleled by the *so doing* construction.

- The *so* in *do so* and preverbal *so* both require linguistic evocation:

- (21) [Andy is holding a newborn baby with one hand behind his head, and shows Gregory]
 Andy: # By *so* holding him, you add support to his developing neck muscles.

- (22) [Andy is holding a newborn baby with one hand behind his head, and shows Gregory]
 Andy: # By *so doing/doing so*, you add support to his developing neck muscles.

In contrast, (postverbal) manner adverbial *so*, like other manner adverbial anaphors, permits situational evocation:

- (23) [Andy is holding a newborn baby with one hand behind his head, and shows Gregory]
 Andy: By holding him *so*, you add support to his developing neck muscles.
- (24) [Andy is holding a newborn baby with one hand behind her head, and shows Gregory]
 Andy: By holding him *this way*, you add support to his developing neck muscles.

- Manner adverbial *so*, like other manner adverbial anaphors, requires an evoked manner:

- (25) a. Andy held the baby in a cradle position. By *holding him so/holding him that way*, he made him stop crying.
 b. Andy held the baby. #By *holding him so/holding him that way*, he made him stop crying.

In contrast, the *so* in *do so* and preverbal *so* do not:

- (26) a. Andy held the baby in a cradle position. By *doing so/so doing*, he made him stop crying.
 b. Andy held the baby. By *doing so/so doing*, he made him stop crying.

Such a characterization of the *do* and the *so* in the *do so* and *so doing* constructions leads us to an analysis in which the *do* denotes the most general type of event, and the *so* marks the information status of that event as discourse-old and salient. Thus, *do so* and *so doing* are simply forms of standard hyponymic reference (cf. Miller 1990).

- (27) John Gotti dispensed with his mob boss by shooting him in broad daylight, with plenty of witnesses around.
- a. By *so shooting him*, Gotti established himself as his victim's likely successor. [same verb]
 b. By *so murdering him*, Gotti established himself as his victim's likely successor. [more general hyponym]
 c. By *so doing*, Gotti established himself as his victim's likely successor. [most general hyponym]

In this way, reference to events with these constructions directly parallels reference to entities with nominal anaphors:

- (28) Chris took his poodle to the vet. The poodle / the dog / the animal was in a lot of pain.

This analogy predicts that *do so* and *so doing* are truly anaphoric in that they are used to refer to events in (the speaker's representation of) the hearer's mental model of the discourse. Specifically, these event anaphors do **not** require an antecedent of any particular syntactic form, and thus are not SURFACE ANAPHORS in the sense of Hankamer and Sag (1976).

naturally-occurring examples of anaphor-antecedent syntactic mismatches:

- Voice mismatches:

- (29) Section 1 provides the examples to be derived by Gapping, and a formulation of Gapping capable of *doing so*. [= deriving the examples] (text of Neijt 1981)
- (30) As an imperial statute the British North America Act could be amended only by the British Parliament, which *did so* on several occasions [= amended an imperial statute]. (=3)

- Nominalized antecedents:

- (31) The defection of the seven moderates, who knew they were incurring the wrath of many colleagues in *doing so*, signaled that it may be harder to sell the GOP message on the crime bill than it was on the stimulus package. [= defecting] (*Washington Post*)
- (32) Even though an Israeli response is justified, I don't think it was in their best interests to *do so* right now. [= respond] (token provided by Dan Hardt)

- Split antecedents:

- (33) The survey results, released by county officials this week, also showed that most of the teenagers who drank alcohol, smoked marijuana or had sex started *doing so* between the ages of 13 and 16. [= drinking alcohol / smoking marijuana / having sex] (*Washington Post*)
- (34) Fortunately, the first person to die in 1990 and the first couple to file for divorce in 1990 were allowed to *do so* anonymously. [= die / file for divorce] (text of Roeper (1990), cited by J. McCawley's "1990 Linguistic Flea Circus" and discussed by Dalrymple, Shieber, and Pereira (1991))

- Other mismatches:

- (35) There was a lot more negativity to dwell on, if anyone wished to *do so*. [= dwell on more negativity]
- (36) With or without the celebration, Belcourt is well worth seeing, and you can *do so* year round. [= see Belcourt]
- (37) ... He went on to claim that the allegedly high-spending Labour authorities had, by *so doing*, damaged industry and lost jobs. [= spent highly] (Cornish 1992)

To summarize thus far, *so doing* and *do so* are both compositional anaphoric constructions consisting of intransitive main verb *do*, denoting the most general of event types, and identifier *so* marking the information status of that event. As predicted, these constructions are not restricted to antecedents of a particular syntactic form.

3 Fu, Roeper, and Borer (2001)

Fu, Roeper, and Borer (2001) (henceforth FRB) assume that *do so* categorically requires a syntactic VP antecedent, and then use this assumption to argue that process nominals such as those in (31–32) contain a VP in their syntactic representations, contra Kehler and Ward (1995). However, there are several serious problems with their analysis.

I. Comparison with ‘deep anaphora’

FRB cite the contrast between (38–39) as evidence for their syntactic antecedent requirement:

- (38) His removal of the garbage in the morning and Sam’s doing so in the afternoon were surprising. (= FRB’s 42b)
- (39) * Kim’s accident in the morning and Sue’s doing so in the evening were not coincidences. (= FRB’s 43b)

However, these judgments do not significantly change when *do so* is replaced by an indisputably deep anaphor such as *do it*, which places no formal restriction on its antecedent:

- (40) His removal of the garbage in the morning and Sam’s doing it in the afternoon were surprising.
- (41) * Kim’s accident in the morning and Sue’s doing it in the evening were not coincidences.

Thus, the contrast cannot be attributed to a requirement that *do so* take a VP antecedent.

II. Non-process nominalizations

FRB’s analysis applies only to process nominalizations, and not (for instance) role nominalizations. But felicitous, naturally-occurring examples of the latter are easily found:

- (42) One study suggests that almost half of young female smokers do so in order to lose weight. [smoke]
- (43) The majority of horse riders do so purely for leisure and pleasure. [ride horses]
- (44) AmericaNet.Com, its officers, directors or employees are not responsible for the content or integrity of any ad. Sellers/buyers/subscribers/investors do so at their own risk. [sell/buy/subscribe/invest]
- (45) Data from the Retirement Survey reveals that 5% of early retirees do so because of the ill health of others. [retire early]

FRB could argue that role nominalizations also have underlying VP structure. But many cases remain unacceptable:

- (46) # My computer does so faster than yours. [compute]
- (47) # The boat’s propeller failed to do so, and now we’re stuck. [propel]

Role nominalizations display gradience with respect to compositionality, which we argue determines the accessibility of the event they nominalize. We know of no **independent** evidence that some role nominals incorporate VP syntax and others do not, nor do we believe that such evidence exists.

III. Other syntactic mismatches

The data that challenge the notion that *do so* requires a syntactic antecedent are not limited to cases involving nominalized antecedents.

CASE 1: TRACE-CONTAINING ANTECEDENTS

Several previously cited examples – (29–30) and (35–37) – involve syntactic mismatches in which the required antecedent VP does not exist. Consider (30), repeated below as (48):

- (48) As an imperial statute the British North America Act could be amended only by the British Parliament, which *did so* on several occasions [= amended an imperial statute].

About this example, FRB say:

it is not clear that it seriously jeopardizes the claim that *do so* requires a VP/V’ antecedent...That in the first conjunct the direct object is occupied by a trace, rather than a full NP may very well turn out to be immaterial for the licensing of the anaphor *do so*. (2001: 572–573)

However, they do not provide the details necessary to evaluate this possibility. Indeed, they never state their assumptions about how *do so* is interpreted nor **why** it would require a syntactic antecedent, especially considering that *do so* is not associated with an ellipsis site.

They contrast such cases which adjectivals, which are presumably worse because they do not involve a trace:

- (49) ?? This act turned out to be amendable, and the British Parliament did so in its last session.

But considerably more acceptable variants are readily constructed:

- (50) After the British Parliament found out that the act was amendable, they elected to do so at their very first opportunity.

CASE 2: SPLIT ANTECEDENTS

FRB also do not address cases in which *do so* is felicitous with a split antecedent – such as examples (33), (34), and (44) – in which a suitable antecedent is not available:

- (51) Fortunately, the first person to die in 1990 and the first couple to file for divorce in 1990 were allowed to *do so* anonymously. (=34)

As such, they fail to capture the similarity with pronominal reference, which is likewise compatible with such referents:

- (52) The first person of the year to die is usually listed in the newspaper and so is the first couple to file for divorce. In a rare show of respect, this year *their* names were kept private.

Lastly, FRB take us to task for failing to capture an alleged parallel between *do so* anaphora and adverbial modification, based on examples such as the following:

- (53) a. Kim's explanation of the problem to the tenants thoroughly (did not prevent a riot). (= FRB's 1a)
b. The occurrence of the accident suddenly (disqualified her). (= FRB's 2a)

However, we have been unable to find a single native informant who finds these examples acceptable (on the relevant readings).

Turning the tables on FRB, we take issue with their failure to capture a different generalization: that the *so* in *do so* is the same (identifier) *so* found in the productive *so+V* construction, which clearly does not require a syntactically-matching antecedent (cf. 13).

4 Conclusions and Future Research

Both preverbal *so* and *do so* require referent events that are:

- hearer-old
- discourse-old
- salient

These forms **do not** impose purely syntactic restrictions on their antecedents, and hence provide no evidence that process nominals incorporate a syntactic VP. On the other hand, there **is** a connection between syntactic form and discourse accessibility, and in turn, discourse accessibility and felicity of reference with *do so*.

We are currently developing an account that identifies and incorporates the determinants of discourse accessibility for different categories of nominalized events, which, like other so-called anaphoric island violations (Ward, Sproat, and McKoon 1991), are both complex and nuanced:

- (54) a. The greatest teachers do so by example. (from our webpage corpus)
b. ?? The tallest teachers do so by example.

References

- Anderson, Stephen R. 1968. Pro-sentential forms and their implications for English sentence structure. Technical Report Report NSF-20, Aiken Computation Laboratory, Harvard University.
- Bolinger, Dwight. 1972. *Degree Words*. Mouton, The Hague, Paris.
- Bouton, Lawrence P. 1970. Do so: Do + adverb. In J. M. Sadock, editor, *Studies Presented to R. B. Lees by his Students*. Linguistics Research, Edmonton, pages 17–38.
- Cornish, Francis. 1992. *So Be It*: The discourse-semantic roles of *so* and *it*. *Journal of Semantics*, 9:163–178.
- Dalrymple, Mary, Stuart M. Shieber, and Fernando Pereira. 1991. Ellipsis and higher-order unification. *Linguistics and Philosophy*, 14:399–452.
- Dechaine, Rose-Marie. 1994. Ellipsis and the position of subjects. In *Proceedings of NELS 24*.
- Fu, Jingqi and Thomas Roeper. 1993. The VP within nominalizations: Evidence from adverbs and the VP anaphor *do-so*. Paper presented at the Syntax-Morphology Connection, MIT.
- Fu, Jingqi, Thomas Roeper, and Hagit Borer. 2001. The VP within process nominals: Evidence from adverbs and the VP anaphor *do-so*. *Natural Language and Linguistic Theory*, 19:549–582.
- Halliday, M.A.K. and Ruqaiya Hasan. 1976. *Cohesion in English*. Longman's, London. English Language Series, Title No. 9.
- Hankamer, Jorge and Ivan Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry*, 7:391–426.
- Kehler, Andrew and Gregory Ward. 1995. On the anaphoric status of *do so*. Paper presented at the Annual Meeting of the Linguistic Society of America, New Orleans, January.
- Lakoff, George and John Robert Ross. 1966. Criterion for verb phrase constituency. Technical Report Report NSF-17, Aiken Computation Laboratory, Harvard University.
- Miller, Philip. 1990. Pseudogapping and *do so* substitution. In Michael Ziolkowski, Manuela Noske, and Karen Deaton, editors, *Papers from the 26th Regional Meeting of the Chicago Linguistic Society*, pages 293–305.
- Neijt, Anneke. 1981. Gaps and remnants – sentence grammar aspects of gapping. *Linguistic Analysis*, 8(1):69–93.
- Prince, Ellen F. 1981. Toward a taxonomy of given-new information. In Peter Cole, editor, *Radical Pragmatics*. Academic Press, New York, New York, pages 223–255.
- Prince, Ellen F. 1992. The ZPG letter: Subjects, definiteness, and information-status. In Sandra Thompson and William Mann, editors, *Discourse Description: Diverse Analyses of a Fundraising Text*. John Benjamins B.V., Philadelphia/Amsterdam, pages 295–325.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Longman Inc., New York.
- Sag, Ivan and Jorge Hankamer. 1984. Toward a theory of anaphoric processing. *Linguistics and Philosophy*, 7:325–345.
- Ward, Gregory, Richard Sproat, and Gail McKoon. 1991. A pragmatic analysis of so-called anaphoric islands. *Language*, 67:439–474.