1 The puzzle: Active by phrases

As is well known, English allows for the agent of a passive sentence to be ‘recovered’ in an optional by phrase, as in (1). What is less frequently discussed is the existence of a parallel construction in active clauses, as shown in (2).

(1) The cake was baked by John.

(2) John baked the cake by himself.

On the surface, these constructions look quite similar; both involve the by preposition in a non-locative function, and both are used to recall the agent of the sentence. In (1), the by phrase provides a home for the displaced/demoted agent, whereas in (2), the meaning of the by phrase is that John carried out the action of baking with no assistance from anyone else (hereinafter, the ‘no help’ reading). Because of its rather obvious function, the passive by phrase is generally set aside for its own unique analysis to the exclusion of other PPs headed by by, for example locatives.

The active by phrase of (2) also has properties which suggest that it warrants a different analysis from the locatives. Most notable is the fact that locatives allow for a flexibility of their complements which the active by phrase does not:

(3) a. John saw a snake beside himself.
   b. John saw a snake beside him.

(4) a. John baked the cake by himself.
   b. #John baked the cake by him.

In (3), the complement of the PP may be either a reflexive pronoun or a referential pronoun with no appreciable difference in meaning. This freedom is widely noted, with Reinhart and Reuland (1993) referring to the type of reflexive usage in (3a) as a logophoric use of a SELF anaphor. Where the free variation between pronominal forms is observed, the key fact is that the
pronominal (reflexive or referential) is not a co-argument of the subject antecedent; such logophoric uses are restricted to non-argument positions.

In (4b), the referential pronoun is only acceptable under a locative reading; i.e. the sentence is saying something about the cake that happens to be in close physical proximity to John. The reading that John did the baking with no outside help is available only with the reflexive. To account for this, either some sort of categorisation of non-argument positions will explain the meaning preservation that exists in (3) but not (4), or a line of reasoning must be pursued to make the claim that there is something argument-like about the ‘no help’ reading. This latter path is the one followed here. Specifically, this paper will argue that the active by phrase has the same argument-like status as the passive by phrase by virtue of the fact that they are one and the same constituent: a vP modifier.

The rest of this paper is organised as follows: section 2 will go into more detail on the usage of the active by phrase, presenting facts drawn from two corpus studies. In section 3, existing analyses of the passive are presented, with an emphasis on claims that have been made about the passive by phrase. Section 4 presents the main proposal of the paper, first presenting meaning-based arguments for the unification of the two by phrases, and then reconciling the syntactic facts from section 3 with the proposed analysis. Section 5 presents a brief conclusion summarising remaining open questions.

2 Distribution of the active by phrase

A valid question in approaching this data may be to wonder whether the active by phrase is worth detailed consideration. While it does present this challenge for binding theory, it may be the case that the construction is so peripheral to the grammar that it does not merit detailed consideration, and can be swept to the sidelines along with other rare (but problematic) data.

Storoshenko (2008) presents a corpus study examining the distribution of reflexive pronouns in English. Among the findings, it is reported that reflexive pronouns within by phrases account for less than 5% of the total instances of reflexive pronouns, with only 50 instances in two million words (one million each of written and spoken text). By and large, these occurred with agentive predicates, never with possessive have or psych-predicates. However, in that study, a handful of examples where the by...-self phrase occurred with copular be were left undiscussed. At best, the picture which emerges is that these phrases are indeed rare, but not entirely well-behaved.

Recently, a new corpus study has been undertaken, with an aim toward learning more about the acquisition of reflexive pronouns. This new study makes a comparison between three sources: “standard” conversational English as drawn from the Corpus of Contemporary American English (Davies 2008), and the CHILDES corpus (MacWhinney 2000). A first pass at these data sources draws 500 tokens randomly from COCA (specifically, transcripts of
Table 1: *by* phrases containing non-logophoric reflexive pronouns across three corpora (out of 500 reflexive tokens per corpus)

<table>
<thead>
<tr>
<th>Phrase type</th>
<th>COCA</th>
<th>Child-Directed</th>
<th>Child-Produced</th>
</tr>
</thead>
<tbody>
<tr>
<td>No help <em>by</em> phrase</td>
<td>4</td>
<td>113</td>
<td>134</td>
</tr>
<tr>
<td>Alone <em>by</em> phrase</td>
<td>3</td>
<td>11</td>
<td>43</td>
</tr>
<tr>
<td>Passive <em>by</em> phrase</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Conversations on Fox News programming, and 500 tokens each from child-directed speech and child speech selected randomly from CHILDES. Improving upon the methodology used in the 2008 corpus study, the present study also makes finer grained distinctions among (obligatory) reflexive pronouns within *by* phrases, identifying three classes as shown in (5):

(5)  a. John baked the cake by *himself*, ('no help' reading)
     b. John went to the store by *himself*, ('alone' reading)
     c. John was frightened by *himself*, (bound passive *by* phrase)

For (5a) and (5b), the context surrounding the *by* phrase was used to make a determination of the meaning. (5c) was included as a separate category as there had been a small number of cases in the 2008 study. Locative uses of *by* were considered for inclusion, but all were judged to be logophoric in the sense of (3), and thus excluded. Counts between the three corpora are given in Table 1.

Dealing first with the distinction between (5a) and (5b), one of the issues outstanding from the previous study has now been resolved. Across all three corpora, it emerged that all those cases where the *by*-...*self* phrase appeared in conjunction with copular *be*, the reading was the 'alone' reading. Making this distinction allowed all the 'no help' readings to be isolated to agentive predicates. (5c), the category for reflexive pronouns in a passive *by* phrase, marked one of the distinctions between the corpora, as such examples only emerged in the adult-to-adult speech.

The more informative contrast is of course in the distinctions the corpora seem to make along the other two uses of the *by*-...*self* phrase. In the adult-to-adult COCA sample, there are only four examples of the ‘no help’ reading, all occurring with transitive predicates for which a passive form could also be derived. However, while these made up less than 1% of the overall reflexive pronouns in the COCA sample, the ‘no help’ *by* phrase made up over 22% of the reflexive pronouns directed toward children, and over 26% of those uttered by children. As with the adult-to-adult cases, all were agentive predicates.

In the passive examples from the COCA sample, both sentences clearly had the reading that the promoted theme in subject position was binding the reflexive:

(6) She was endlessly surprised by herself.
Though rare, such examples do emerge. Where they do, the reading appears to be constrained by the animacy of the promoted object. In (6), there is a clear reading where the subject is self-surprising; the same is not true for (7):

(7) The cake, was baked by itself.

This sentence cannot have the meaning that the cake was self-baking; at best it will have the meaning that there was nothing else in the oven while the baking was taking place.

The ‘no help’ reading does not seem to be available at all for passives. Certainly, adding an unbound reflexive does not result in anything grammatical:

(8) *The cake was baked by himself.

Some speakers report that having two by phrases on a passive (one a passive by phrase, and the other contributing the ‘no help’ meaning) may be possible, but judgements are mixed, and no such examples are found in the corpus:

(9) ?The cake was baked by John, by himself.

To get the reading where John did the baking with no help, it seems that an active must be used. The incompatibility of the ‘no help’ reading of the by…self phrase with the passive is the first piece of evidence that the active and passive by phrases should be unified. (7) makes it clear that the ‘alone’ reading of the by…self phrase comes though with no problem in the passive, motivating the distinction in the corpus study, while the contrast between (6) and (7) reaffirms the connection between the ‘no help’ reading and conscious agency.

Returning to the facts from the corpus study, the preponderance of by…self phrases with the ‘no help’ meaning in the CHILDES samples demonstrates that these phrases appear to play some sort of important role in the acquisition of Condition A. While adults speaking with each other tend to use reflexive pronouns in Condition A compliant argument positions only about half the time, cases like (5a) provide children with an additional set of examples where reflexive pronouns are obligatory and tend to follow Condition A. Based on this fact alone, it seems as though arriving at a firm analysis of the ‘no help’ by…-self phrase is a worthwhile endeavour, as it forms a disproportionately large portion of the total data available for children acquiring Condition A.

3 Treatments of the passive by phrase

Having established both the importance of the ‘no help’ meaning by…self phrase and its incompatibility with passives, here more details on the analysis of
the passive are discussed, with an aim toward finding common facts about passive and active by phrases. First, the classic Jaeggli (1986) analysis of the passive is reviewed, followed by two more recent approaches.

3.1 The classic analysis

The predominant analysis of the passive finds its roots in Jaeggli (1986). This is the familiar story in which the passive morphology absorbs both the external theta role of a predicate and the accusative case assignment. Lacking case, the internal argument moves to the subject position, and the external theta role is optionally assigned to an adjunct by phrase via the verbal projection:

(10)

Jaeggli has little to say on the subject of the active by...-self phrases, only making the claim that in active clauses by phrases may only be instrumental or locative. The ‘no help’ by...-self phrase is not mentioned at all.

Aside from this oversight, the lack of any real formal definition of the transmission mechanism for the external theta role has remained a weakness of this analysis. To resolve this issue, two more recent approaches to the passive take the position that the external (or agent) theta role is uniformly assigned in both actives and passives.

3.2 The smuggling analysis

Collins (2005) presents a novel analysis of the derivation of the passive in which external theta roles are uniformly assigned in [Spec, vP]. The preposition by in
the passive is re-analysed as the head of VoiceP, and movement of the whole Part(icle) Phrase to a position left of Voice', followed by further extraction of the internal argument yields the canonical passive word order:

(11)

To motivate this analysis, Collins makes the following observations about the position of the passive by phrase: it cannot intervene with a particle verb (12), nor can it appear to the right of a preposition stranded in a pseudopassive (13):

(12)
   a. *The argument was summed by the coach up.
   b. *The paper was written by John up.

(13)
   a. *John was spoken by Mary to.
   b. *The clown was laughed by the children at.

These facts are taken as evidence for the PartP movement which is key to this analysis. However, it is worth noting that the ‘no help’ by...-self phrase has exactly the same restrictions:

(14)
   a. *The coach, summed the argument by himself, up.
   b. *John, wrote the paper by himself, up.

(15)
   b. *The children, laughed by themselves, at the clown.
As such, the facts actually point toward an analysis under which the passive by phrase and the ‘no help’ by...-self phrase are in fact the same constituent.

In cases where there is no by phrase, Collins posits a covert agent in [Spec, vP], to which the external theta role is assigned. To motivate this, he cites binding facts as in (16):

(16)a. Such privileges should be kept to oneself.
    b. Damaging testimony is always given about oneself in secret trials.

The reflexive pronouns in (16) are argued to be bound by the covert external argument in [Spec, vP]. For this to be the case, it is worth noting that the relevant binding relation must obtain at a point in the derivation before PartP moves out of the c-command domain of the external argument position. For binding theoretic purposes then, even passive sentences are considered at a point where the external argument c-commands the internal arguments. As such, Collins makes a clear prediction that reflexive pronouns should never occur inside a by phrase with a co-indexed theme; such a reflexive would itself be unbound, and it would in turn trigger a Condition B or C violation, locally binding some other nominal. Naturally occurring examples such as (6) though, where the reflexive is indeed in a passive by phrase do crop up in both the COCA corpus extract and the 2008 study. That such examples are grammatical casts doubt on Collins’ analysis, and remains a fact which will need to be captured in the final analysis.

3.3 Universal order of merge

A second attempt to unify the theta role assignment between actives and passives comes in Bowers (2010). This approach departs from both Jaeggli and Collins in that agent theta roles are not assigned in [Spec, vP], but rather as the specifier of an Agent-selecting head. Bowers proposes a series of functional heads selecting various thematic roles, and provides a fixed universal order of merge for these heads. These heads are merged above a verbal root, with each successive head embedding those coming before. Head movement and EPP/case-driven movements are employed to derive word order. Elements introduced in PPs are fixed in position, ineligible for movement. The ordering of the heads relevant to the present discussion is given in (17), yielding the schematic structure in (18):

(17) \text{Place} < \text{Manner} < \text{Ag(ent)} < \text{Th(eme)}
A crucial fact about passives in this analysis is that a post-verbal by phrase reflects the base-generation position for the agent, c-commanded by the theme. While this will capture the facts about the binding of passive by phrases missed under Collins’ analysis, this analysis makes some clear predictions under the assumption that the ‘no help’ by…-self phrase should be introduced under the manner functional head.

First, this predicts the possibility of marginal examples such as (9), in which the passive by phrase and the ‘no help’ by…-self phrase were constructed into the same example. Recall also that (8) employed a by phrase with the ‘alone’ reading that was acceptable in the passive. This reading is even compatible with the passive by phrase:

(19a) The cake was baked by itself by John.
(19b) ?The cake was baked by John by itself.

(19a) presents the most natural order for the sentence, while (19b) presents the order predicted assuming the tree structure in (18). Similar facts obtain with passives and locatives PPs:

(20a) A snake was seen beside himself by John.
(20b) ?A snake was seen by John beside himself.

In these cases, the passive by phrase should be out at the right periphery, contrary to the structure in (18). At the very least, these facts cast some doubt on the analysis.
4 Unifying the by phrases

In the prior section, two approaches to unifying the active and passive external arguments were shown to yield ultimately untenable positions. Based particularly on the facts observed in the analysis of Collins’ argumentation, it seems that a different unification is motivated. Rather than assuming that the external argument is assigned in the same way in both the active and the passive, it should be the case that the two by phrases in (21) are one and the same:

(21) a. John, baked the cake by himself.
   b. The cake was baked by John.

Not only will this capture the distributional facts already discussed, but it represents a novel way of addressing the issue of external theta role assignment from Jaeggli. In (21a), the by…-self phrase cannot possibly be the recipient of the external theta role, that having been assigned to John. By extension, assuming the passive by phrase to be one and the same constituent, the issue of assigning the external theta role is neatly sidestepped: this is simply not a constituent to which that role need be assigned. In the balance of this section, the proposed treatment of the by phrase will be presented.

4.1 A common semantics

The meaning of the ‘no help’ by…-self phrase is clear enough; affirming that the agent carried out the described action with no additional assistance. In (22), it is shown that partial co-reference for this by phrase is not acceptable:

(22) a. *John baked the cake by themselves.
   b. *John and I baked the cake by myself.

In other words, this phrase affirms that the described event has no other co-agents. Though not so immediately apparent, the same exclusivity is carried by the passive by phrase, which also resists partial co-reference in either direction:

   #The cake was baked by us.
   b. Context: John and Mary baked the cake together.
   #The cake was baked by John.

Just as in the active, the sentences of (23), placed in an explicit context, show that there must be a one-to-one correspondence between the complement of the passive by phrase and the agent of the described event, even though that agent is not otherwise represented in the sentence.
The semantics of this *by* phrase in both active and passive is thus to assert that the complement of *by* is the only agent participating in the described event. This can be expressed in a neo-Davidsonian event semantics using a form similar to the semantics of *only*. (24) gives the semantic form for an unadorned active transitive, followed by both an active and a passive with the added *by* phrase.

(24) a. John baked the cake.
   \[\exists e [BAKE(e) \land AGENT(e)=\text{john} \land THEME(e) = \text{cake}]\]

b. John baked the cake by himself.
   \[\exists e [BAKE(e) \land AGENT(e)=\text{john} \land THEME(e) = \text{cake} \land [AGENT(e)=y \rightarrow y = \text{john}]]\]

c. The cake was baked by John.
   \[\exists e [BAKE(e) \land THEME(e) = \text{cake} \land [AGENT(e)=y \rightarrow y = \text{john}]]\]

As shown, the formation of the simple event in (24a) defines John as agent of the event, but does not exclude the possibility of other co-agents. This meaning is added in (24b), and the same meaning again appears in (24c), though non-redundantly as the agent is not otherwise specified. Note that the semantic contribution of the *by* phrase does not directly express the agent of the event, but rather a condition on agents; in (24b), a complement referring to anyone other than *John* would result in a false statement and thus a crashed derivation.

4.2 Capturing the syntax

Given that this *by* phrase is exclusively tied to the external theta role of a predicate, the most logical position will be to treat it as an adjunct to *vP*. This assumes that the *v* head itself is the origin of that theta role, and that the role is available locally throughout the *vP* projection. Overtly projected to [Spec, *vP*] in actives but not passives, this role is merely referenced in the semantics of the *by* phrase modifier. Indeed, the difference between active and passive *v* may reduce down to whether or not this role is obligatorily expressed through the specifier position; optional (indirect) expression via the modifier is always available.

To capture the observed binding facts, two relatively non-controversial further assumptions must be made. First is the Chomsky (2005) claim that active *vP* is a phase boundary while passive *vP* is not. As a phase boundary, the specifier of *vP* should be the “escape hatch” allowing upward movement; this translates to that specifier being the highest element of the *vP*. The *by* phrase modifier will remain within the *vP*, but crucially below that highest specifier, the base-generation position of the external argument in actives. Consequently, the external argument of an active *e*-commands the *by* phrase. This is shown in the tree structure in (25).
The second assumption derives from Canac-Marquis (2005), who claims that phase boundaries are binding domains. By virtue of the semantics, the by phrase complement is already obligatorily co-referential with the external argument; these syntactic assumptions combine that co-reference with c-command within the relevant binding domain. Naturally then, the reflexive pronoun is obligatory.

In passives, the situation is somewhat different. With no [Spec, vP], the derived subject is the only nominal c-commanding the by phrase. Furthermore, as the passive vP is not a phase boundary, it is not a binding domain. Rather, the whole clause becomes the binding domain for the by phrase complement. As such, one again expects an obligatory reflexive pronoun in those cases where the promoted internal argument happens to be co-referential with the by phrase complement. This scenario is illustrated in (26), paraphrasing the corpus example from (6):
This yields the correct prediction, as only a reflexive pronoun is possible where the internal argument and the by phrase complement are coreferential. While such examples are degraded, an analysis predicting that only a reflexive pronoun is possible in this position is on the right track. The rarity of such examples is ascribed not to the grammar, but rather to a stylistic choice between synonymous forms. Thus, the simplicity of the Jaeggli-style passive is preserved while accounting for earlier problematic data. Note also that by placing the adjunct relatively high in the right periphery (above the usual positions for locative and manner adjuncts), the facts in (19) and (20) are accounted for.

5 Conclusion

In this paper, corpus evidence has been cited to make the case that the ‘no help’ by...-self phrase plays an important role in the acquisition of Condition A, and warrants a more rigorously-defined syntactic treatment. Furthermore, by looking at passive by phrases, it emerges that there is enough common ground between to warrant a unification of the two. In so doing, examples which were troublesome when the two phenomena were treated separately are accounted for, and a lingering doubt surrounding treatments of the passive has been addressed without the need for any wholesale re-imagining of the passive derivation in English. In further work, it remains to be seen whether the acquisition of passive by phrases tracks with the use of the ‘no help’ by phrase, as predicted under this analysis.

References


Dennis Ryan Storoshenko
Yale University
dennis.storoshenko@yale.edu