English VP anaphors: *do it, do this, do that*

Gabriel Flambard

Université Paris Diderot
ED 132 – Sciences du Langage
CLILLAC-ARP

Ellipsis Day
July 9th, 2014
**VPAs in English**

**VP anaphora**

- *Do it*, *do this* and *do that* may all refer to an action antecedent expressed by a VP
- In this usage, alternate with *do so* and PAE

(1) a. With the move to the new system, we’re abolishing *that subsidy*, but we’re **doing it** in two goes: 50% of it this year and all of it will go next year. (KRT)

b. We’re **doing this/that** in two goes...

(2) We’re **doing so** in two goes

(3) We’re going to abolish *that subsidy*, and when we **do**, a new system will be put into place
VPAs: the alternation puzzle

VPA choice

- What drives the choice of a particular VPA in context?
- Few studies on this issue, but see Huddleston and Pullum (2002); Miller (2011) for preliminary accounts

A question of choice...

(4) We’re abolishing that subsidy, but we’re doing it (ok: doing this/that) in two goes: 50% of it this year and all of it will go next year. (KRT)

Or no choice:

(5) He closes his eyes when he speaks and I don’t trust anyone who does that. (does this/#it). (AHF)
VPAs: the alternation puzzle

- VPA alternation is relatively free in some contexts and more restricted in others.
- Minor differences can ‘change the deal’ in terms of alternation possibilities:

(6) He closes his eyes when he speaks and I don’t trust anyone who does that. (does ?this/#it). (AHF)

(7) He closes his eyes when he speaks. I wonder why he does that/this/it.
Antecedent: a clarification

Following Cornish (1999), it is useful to distinguish between

- **Antecedent** in the *conceptual* sense, a representation of the anaphor’s referent in a discourse model
- **Antecedent** in the *structural* sense: the structure whereby an antecedent is accessed, Cornish’s *antecedent-trigger*
Aims of the talk

- Investigate VPA alternation based on an annotated corpus
- Focus on methods: coding strategies, tendencies emerging from the data
- Identify the relevant factors of choice
The corpus

900 occurrences from the BNC:

- 500 *do it*; 200 each *do this/that*
- Source: written section (fiction) and various spoken texts (news reports, TV/radio debates...)

**Features coded:** 12 in total, these include:

- **Syntactic** Finiteness, presence of an adjunct, distance from antecedent-trigger, structure of the trigger/VPA
- **Semantic** Type of adjunct, identity of subjects/actions between the VPA and trigger
- **Discourse** Topicality of antecedent: previous discourse topic, or continued topic beyond the VPA sentence

Coding still in progress, results preliminary
Outline

We leave out topicality and consider:

1. **Finiteness**: finite or non-finite *do*
2. **Adjuncts**: presence or absence, type of adjunct
3. **Identity of subjects**: whether the VPA has the same subject as its antecedent-trigger
4. **Identity of SoAs**: whether the VPA and antecedent-trigger denote the same state of affairs
5. **Structure** of the antecedent-trigger and VPA clause
Finite and non-finite VPAs

Known facts:

- VPAs are mostly **non-finite**, see Houser (2010); Miller (2013) for *do so*. Present data confirm the trend for *do it/this/that*.
- Stative *do so* more common in non-finite uses, since PAE is not available (see Huddleston and Pullum, 2002)

(8) Landsburg attempts (...) to argue that racism does not **exist** in corporate America because it would not be in its bottom-line interest for it to **do so**. (Houser, ex. 1d)

- No such interaction with *do it/this/that*: states are bad even in non-finite uses
- Influence of finiteness elsewhere is an open question
Finite and non-finite VPAs

Finite and non–finite uses of VPAs

Do it

Do this

Do that

References

English VP anaphors: do it, do this, do that

Gabriel Flambard

VP anaphora: overview

Finiteness

Adjuncts

Presence or absence of an adjunct

Adjunct types

Identity of subjects and states of affairs

Identity of subjects

Identity of states of affairs

Structure

Structure of the antecedent-trigger

Structure of the VPA

References
English VP anaphors

1. VP anaphora: overview

2. Finiteness

3. Adjuncts
   - Presence or absence of an adjunct
   - Adjunct types

4. Identity of subjects and states of affairs
   - Identity of subjects
   - Identity of states of affairs

5. Structure
   - Structure of the antecedent-trigger
   - Structure of the VPA
Adjuncts and VP As

- Finite *do so* strongly prefers having an adjunct (83% of cases in Miller (2011))
- *Do so* and *do it* preferred over PAE if an adjunct is present (Houser, 2010; Miller and Hemforth, 2014)
- True mainly if the adjunct is **non-contrastive**: simply specifies an additional property of the antecedent

(9) A.—Does he *shop in women’s*?
B.—He *does it/#does* to find things his size. (M&H, 4b)
Adjuncts: frequency

Frequency of adjuncts by VPA

Adjunct types

Do it

Do this

Do that
Adjuncts and finiteness

**Adjuncts with finite/non-finite VPAs**

<table>
<thead>
<tr>
<th>Adjunct</th>
<th>finite</th>
<th>non-finite</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do it</td>
<td>60%</td>
<td>40%</td>
</tr>
<tr>
<td>Do this</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>Do that</td>
<td>70%</td>
<td>30%</td>
</tr>
</tbody>
</table>

- Adjunct: Presence of an adjunct
- No adjunct: Absence of an adjunct
Data summary

Overall:

- **Do it**: no preference
- **Do this/that**: mostly no adjunct, especially *do that*

Variable influence of finiteness on use of adjuncts:

- **Do it**: adjuncts more frequent in finite uses; nonfinite: no clear preference
- **Do this**: preference for +Adj in finite uses, −Adj in non-finite uses
- **Do that**: preference for −Adj constant in finite/non-finite uses
English VP anaphors

1. VP anaphora: overview
2. Finiteness
3. Adjuncts
   - Presence or absence of an adjunct
   - Adjunct types
4. Identity of subjects and states of affairs
   - Identity of subjects
   - Identity of states of affairs
5. Structure
   - Structure of the antecedent-trigger
   - Structure of the VPA
Adjunct types

- Semantic type of the adjunct: place, time, manner, etc.
- Current coding does not cover cases with more than one adjunct: only the adjunct immediately after the VPA is considered
Adjunct types: *do it*

![Adjunct types: Do it](chart.png)
Adjunct types: *do this*

<table>
<thead>
<tr>
<th>Adjunct type</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>manner</td>
<td>30</td>
</tr>
<tr>
<td>frequency</td>
<td>20</td>
</tr>
<tr>
<td>remnant</td>
<td>10</td>
</tr>
<tr>
<td>time</td>
<td>0</td>
</tr>
<tr>
<td>beneficiary</td>
<td>10</td>
</tr>
<tr>
<td>duration</td>
<td>0</td>
</tr>
<tr>
<td>purpose</td>
<td>0</td>
</tr>
</tbody>
</table>

English VP anaphors: *do it, do this, do that*

Gabriel Flambard

VP anaphora: overview

Finiteness

Adjuncts

Presence or absence of an adjunct

Adjunct types

Identity of subjects and states of affairs

Identity of subjects

Identity of states of affairs

Structure

Structure of the antecedent-trigger

Structure of the VPA

References
Adjunct types: *do that*

Adjunct types: Do that

<table>
<thead>
<tr>
<th>Adjunct type</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td>30</td>
</tr>
<tr>
<td>remnant</td>
<td>20</td>
</tr>
<tr>
<td>place</td>
<td>10</td>
</tr>
<tr>
<td>duration</td>
<td>10</td>
</tr>
<tr>
<td>frequency</td>
<td>10</td>
</tr>
<tr>
<td>manner</td>
<td>10</td>
</tr>
</tbody>
</table>

English VP anaphors: *do it, do this, do that*

Gabriel Flambard

VP anaphora: overview

Finiteness

Adjuncts

Presence or absence of an adjunct

Adjunct types

Identity of subjects and states of affairs

Identity of subjects

Identity of states of affairs

Structure

Structure of the antecedent-trigger

Structure of the VPA

References
Adjunct types: discussion

Frequent adjunct types seem to reflect broader semantic preferences:

- *Do it* typical with manner expressions, including adjuncts and manner interrogative (e.g. *how to do it*, 13 cases) or nouns (e.g. *a way to do it*, 11 cases)

- Unlike *do that*: 3 manner adjuncts, 1 *the way to do that*

Summary: adjuncts and adjunct types

- Distribution of adjunct types across VPAs hints at differences in their discourse properties

- Not used with adjuncts equally often, and not with the same types of adjuncts

- Broader question: how to explain the observed variations in usage and type if adjuncts?
English VP anaphors

1. VP anaphora: overview
2. Finiteness
3. Adjuncts
   - Presence or absence of an adjunct
   - Adjunct types
4. Identity of subjects and states of affairs
   - Identity of subjects
   - Identity of states of affairs
5. Structure
   - Structure of the antecedent-trigger
   - Structure of the VPA
Identity of subjects

Whether or not the VPA has the same subject as the antecedent-trigger (if there is one)

Mostly straightforward, except for:

- Null subjects: identity based on the interpretation of the null subject (10)
- Passive VPs: identity determined by the passive agent (explicit or implicit) (11)

(10) [O]ne would have to think in terms of either publishing the whole thing exactly as it stood, or not doing it at all. (A08, same subjects)

(11) I mean, is that an attitude that should be changed? And if so, how do you do it? (FLD, same subjs)
Identity of subjects

Identity of subjects in the VPA and antecedent–trigger

<table>
<thead>
<tr>
<th></th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>do it</td>
<td></td>
</tr>
<tr>
<td>do this</td>
<td></td>
</tr>
<tr>
<td>do that</td>
<td></td>
</tr>
</tbody>
</table>

- **same**
- **different**
- **N_A**

Gabriel Flambard

References
Identity of subjects: discussion

Clear preference for same subjects for all VPAs

- Also the norm across sentences: next S typically continues the topic of previous S, and has the same subject
- However, VPAs don’t always occur in the S immediately following the antecedent-trigger

Influence of finiteness:

- Same-subject pattern dominant in finite/nonfinite uses
- Some minor differences with *do this/that*, none with *do it*
English VP anaphors

1. VP anaphora: overview

2. Finiteness

3. Adjuncts
   - Presence or absence of an adjunct
   - Adjunct types

4. Identity of subjects and states of affairs
   - Identity of subjects
   - Identity of states of affairs

5. Structure
   - Structure of the antecedent-trigger
   - Structure of the VPA
Identity of SoAs

Whether the VPA denotes the same specific action (SoA) as the antecedent-trigger, or merely the same kind of action.

That is to say:

- The exact same situation (same time and participants), or
- A different of occurrence of the same kind of situation (different time/place, other participants)

(12) I am going to save the business, Joey, I don’t know how I’m going to do it but somehow I will. (CKD)

(13) Anyway, so what if there were people smoking a little—everybody does it, the teachers in school do it...(APU)
Identity of SoAs

English VP anaphors: do it, do this, do that

Gabriel Flambard

VP anaphora: overview

Finiteness

Adjuncts

Presence or absence of an adjunct
Adjunct types

Identity of subjects and states of affairs

Identity of subjects
Identity of states of affairs

Structure

Structure of the antecedent-trigger
Structure of the VPA

References

<table>
<thead>
<tr>
<th>0</th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>do it</td>
<td></td>
<td></td>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>do this</td>
<td>50</td>
<td></td>
<td></td>
<td>20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>do that</td>
<td></td>
<td>40</td>
<td></td>
<td></td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Identiy of SoAs by VPA

- same
- different
- N_A
Identity of SoAs: discussion

Main observations:

- Same SoAs are the norm for all VPAs
- Likely in part due to the preference for identical subjects

Here too, little effect of finiteness:

- Preference for same SoAs in both finite and non-finite uses
- Proportion of identical SAs does not vary with finiteness
English VP anaphors

1. VP anaphora: overview

2. Finiteness

3. Adjuncts
   - Presence or absence of an adjunct
   - Adjunct types

4. Identity of subjects and states of affairs
   - Identity of subjects
   - Identity of states of affairs

5. Structure
   - Structure of the antecedent-trigger
   - Structure of the VPA

References
Structure: coding methods

Structure of the VPA/trigger coded at two levels:

- The structure containing the VPA or its antecedent-trigger, e.g. main VP, complement of a V/N...
- The nature of the clause containing the VPA/trigger: main/subordinate and declarative, interrogative...
Main categories distinguished (shown here for the VPA):

- **Main VP** of the sentence, e.g. *I did it*
- **Modal** Under the scope of a modal, e.g. *I might do it*
- **Complement** of a verb, noun, or adjective, e.g. *I try doing it*; *a way to do it*; *able to do it*

The trigger may also be one of the following:

- **passive VP** As opposed to active
  - **NP** Rather than a VP
  - **VPA** Another occurrence of a VPA
Trigger structure: *do it*

![Graph showing the frequency of different elements in the trigger structure of "do it".]
Trigger structure: *do this*

English VP anaphors: *do it, do this, do that*

Gabriel Flambard

VP anaphora: overview

Finiteness

Adjuncts
- Presence or absence of an adjunct
- Adjunct types

Identity of subjects and states of affairs
- Identity of subjects
- Identity of states of affairs

Structure
- Structure of the antecedent-trigger
- Structure of the VPA

References
Trigger structure: *do that*

Do that

<table>
<thead>
<tr>
<th>Structure</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>main_VP</td>
<td>25</td>
</tr>
<tr>
<td>Vcomp</td>
<td>15</td>
</tr>
<tr>
<td>modal</td>
<td>10</td>
</tr>
<tr>
<td>N_A</td>
<td>5</td>
</tr>
<tr>
<td>VPA</td>
<td>0</td>
</tr>
</tbody>
</table>

English VP anaphors: *do it, do this, do that*

Finiteness

Adjuncts

Presence or absence of an adjunct

Adjunct types

Identity of subjects and states of affairs

Identity of subjects

Identity of states of affairs

Structure

Structure of the antecedent-trigger

Structure of the VPA

References
Triggger structure: discussion

**Overall tendencies**

- Fairly similar distribution of structure types across VPAs
- Main VPs are the most common structures

**But some interesting differences**

- Main VP more common with *do this/that*
- *Do it:* most frequent form to have a VPA as its antecedent-trigger (it is in most cases *do it*. Much rarer with *do this/that*.
- *Do this:* highest proportion of cases where no antecedent-trigger is present (many exophoric uses)
The nature of the trigger clause, coded according to:

- Syntactic type, e.g. declarative vs interrogative
- Main or subordinate status (subordinates include complements: Vcomp etc)
- Relatives, indirect interrogatives, and (pseudo)-clefts marked as such
Trigger Clause type: *do it*

<table>
<thead>
<tr>
<th>Clause type</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>subordinate</td>
<td>40</td>
</tr>
<tr>
<td>main (DECL.)</td>
<td>30</td>
</tr>
<tr>
<td>NA</td>
<td>10</td>
</tr>
<tr>
<td>relative</td>
<td>5</td>
</tr>
<tr>
<td>dir. interrogative</td>
<td>5</td>
</tr>
</tbody>
</table>

English VP anaphors: *do it, do this, do that*

Gabriel Flambard

VP anaphora: overview

Finiteness

Adjuncts
Presence or absence of an adjunct
Adjunct types

Identity of subjects and states of affairs
Identity of subjects
Identity of states of affairs

Structure
Structure of the antecedent-trigger
Structure of the VPA

References
Trigger Clause type: *do this*

<table>
<thead>
<tr>
<th>Clause type</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>subordinate</td>
<td>40</td>
</tr>
<tr>
<td>main (decl.)</td>
<td>20</td>
</tr>
<tr>
<td>NA</td>
<td>10</td>
</tr>
<tr>
<td>relative</td>
<td>5</td>
</tr>
<tr>
<td>dir. interrogative</td>
<td>0</td>
</tr>
</tbody>
</table>
Trigger Clause type: *do that*

![Graph showing frequency of different clause types for *do that*]

- Subordinate
- Main (declarative)
- NA
- Directive interrogative
- If-clause

<table>
<thead>
<tr>
<th>Clause Type</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subordinate</td>
<td>40</td>
</tr>
<tr>
<td>Main (declarative)</td>
<td>54</td>
</tr>
<tr>
<td>NA</td>
<td>6</td>
</tr>
<tr>
<td>Directive interrogative</td>
<td>4</td>
</tr>
<tr>
<td>If-clause</td>
<td>6</td>
</tr>
</tbody>
</table>

References

English VP anaphors: *do it, do this, do that*

Gabriel Flambard

VP anaphora: overview

Finiteness

Adjuncts

Presence or absence of an adjunct

Adjunct types

Identity of subjects and states of affairs

Identity of subjects

Identity of states of affairs

Structure

Structure of the antecedent-trigger

Structure of the VPA

References
Trigger Clause type: discussion

- Subordinates and declarative main clauses are the most common types.
- No notable differences between VPAs apart from the high frequency of cases where *do this* has no antecedent-trigger.
English VP anaphors

1. VP anaphora: overview
2. Finiteness
3. Adjuncts
   - Presence or absence of an adjunct
   - Adjunct types
4. Identity of subjects and states of affairs
   - Identity of subjects
   - Identity of states of affairs
5. Structure
   - Structure of the antecedent-trigger
   - Structure of the VPA
Coding methods

- Same as with the antecedent-trigger before: containing structure and type of clause (but structure is always a VP, and no ‘NA’ cases)
- Only active uses of VPAs are considered: no passives in the corpus
English VP anaphors: *do it, do this, do that*

Gabriel Flambard

VP anaphora: overview

Finiteness

Adjuncts

Presence or absence of an adjunct

Adjunct types

Identity of subjects and states of affairs

Identity of subjects

Identity of states of affairs

Structure

Structure of the antecedent-trigger

Structure of the VPA

References
VPA structure: *do this*

![Diagram showing the structure of VPA with labels for main_VP, Vcomp, modal, Adj_comp, Ncomp, and gerund. The frequency (%) is indicated along the y-axis, ranging from 0 to 50. The main_VP is the most frequent component, followed by Vcomp, modal, and Adj_comp. Ncomp and gerund are the least frequent.]
VPA structure: *do that*

VPA structure: Do that

- Frequency (%)
  - 0
  - 10
  - 20
  - 30
  - 40
  - 50

**main_VP**

**modal**

**Vcomp**

**Ncomp**

**Adj_comp**

**gerund**

References

Structure

Structure of the antecedent-trigger

VP anaphora: overview

Finiteness

Adjuncts

Presence or absence of an adjunct

Adjunct types

Identity of subjects and states of affairs

Identity of subjects

Identity of states of affairs

Gabriel Flambard

English VP anaphors: *do it, do this, do that*
Main VPs dominate (as with trigger structure), but modals are more common

But preference for these structures is stronger than in the antecedent-trigger

Limited differences between VPAs, exc. higher frequency of modals with do it
VPA clause type: *do it*

English VP anaphors: *do it, do this, do that*

Gabriel Flambard

VP anaphora: overview

Finiteness

Adjuncts
Presence or absence of an adjunct
Adjunct types

Identity of subjects and states of affairs
Identity of subjects
Identity of states of affairs

Structure
Structure of the antecedent-trigger
Structure of the VPA

References
VPA clause type: *do this*

- Subordinate
- Main (declarative)
- Direct interrogative
- Relative
- If-clause

<table>
<thead>
<tr>
<th>Clause type</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subordinate</td>
<td>50</td>
</tr>
<tr>
<td>Main (declarative)</td>
<td>20</td>
</tr>
<tr>
<td>Direct interrogative</td>
<td>10</td>
</tr>
<tr>
<td>Relative</td>
<td>5</td>
</tr>
<tr>
<td>If-clause</td>
<td>5</td>
</tr>
</tbody>
</table>

English VP anaphors: *do it, do this, do that*

Gabriel Flambard

Finiteness

Adjuncts
- Presence or absence of an adjunct
- Adjunct types

Identity of subjects and states of affairs
- Identity of subjects
- Identity of states of affairs

Structure
- Structure of the antecedent-trigger
- Structure of the VPA

References
VPA clause type: *do that*

<table>
<thead>
<tr>
<th>Clause type</th>
<th>Frequency (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>subordinate</td>
<td>50 / 54</td>
</tr>
<tr>
<td>main (decl.)</td>
<td></td>
</tr>
<tr>
<td>dir. interrogative</td>
<td></td>
</tr>
<tr>
<td>if–clause</td>
<td></td>
</tr>
<tr>
<td>relative</td>
<td></td>
</tr>
</tbody>
</table>

English VP anaphors: *do it, do this, do that*

Gabriel Flambard

VP anaphora: overview

Finiteness

Adjuncts

Presence or absence of an adjunct

Adjunct types

Identity of subjects and states of affairs

Identity of subjects

Identity of states of affairs

Structure

Structure of the antecedent-trigger

Structure of the VPA

References
Discussion

Subordinate or main status is important:

- *Do it*: typically main clause
- *do this*/that even more clearly prefer subordinates
- True also if we include S types coded separately: relatives, interrogatives etc.

Less so syntactic type:

- No preference for e.g. interrogative vs imperative or relative
- No variation across VPs in this respect
- But, very few tokens, acceptability may vary
Conclusions: factors of VPA alternation

Factors apparently influencing VPA choice:

- **Adjuncts**: use of adjuncts as well as semantic type
- **Main/subordinate** nature of the VPA clause
- Structure of the antecedent-trigger, in particular VPAs (*do it*) or absence of a trigger (*do this*)

Other factors considered are not a source of variation, relevant to the use of each form rather than choice:

- Identity of subjects/states of affairs
- Clause type of the antecedent-trigger
- Structure of the VPA and syntactic type of clause


