Miller, Philip & Hemforth, Barbara. 2024. Verb Phrase Ellipsis with nominal antecedents: the case of polar nouns. *Glossa: a journal of general linguistics* 9(1). pp. 1–43. DOI: https://doi.org/10.16995/glossa.15303

Open Library of Humanities

Verb Phrase Ellipsis with nominal antecedents: the case of polar nouns

Philip Miller, Université Paris Cité, FR, philip.miller@u-paris.fr

Barbara Hemforth, CNRS-Université Paris Cité, FR, barbara.hemforth@u-paris.fr

This paper brings new theoretical and experimental evidence to bear on the classical question of whether Verb Phrase Ellipsis requires a syntactically identical antecedent and how it differs, in this respect, from the Verb Phrase Anaphor *do it*. We focus on the case of VPE with nominal antecedents and argue that it is always grammatical. Its acceptability is variable and is predicted by a discourse condition on VPE, namely that VPE prefers a salient alternative in the discourse context and that the elliptical clause must select one branch of the alternative (Miller & Pullum 2014). We argue that nominal antecedents are only acceptable in those special cases where a noun can express such an alternative. Specifically we focus on the case of 'polar nouns', which can function as a type of polar concealed question, making a polar alternative salient. We provide experimental evidence, based on online acceptability judgments, showing (i) that VPE is judged more acceptable than *do it* when the discourse condition is met, regardless of whether the alternative is expressed by a verb or a noun; (ii) that nominal antecedents are judged slightly less acceptable than verbal antecedents in such cases. In conclusion, we argue that VPE simply requires accessing an antecedent satisfying the discourse constraints within the context but that the heuristic strategies of the parser make use of all available evidence, including syntactic structure in short term memory, making syntactically identical antecedents easier to process and hence slightly more acceptable.

1 Introduction

Since the earliest discussions of Verb Phrase Ellipsis (VPE) in generative grammar, there has been a debate as to whether it is licensed by syntactic identity of the ellipted material with the antecedent or whether it simply requires that the antecedent be recoverable from the discourse context (see among many others, Hankamer & Sag 1976; Sag 1976; Schachter 1977; Sag & Hankamer 1984; Dalrymple et al. 1991; Hardt 1993; Fiengo & May 1994; Johnson 2001; Merchant 2001; Kehler 2002; Dalrymple 2005; Merchant 2008; Kertz 2010; Culicover & Jackendoff 2012; Kertz 2013; Merchant 2013; Frazier & Duff 2019; Poppels & Kehler 2018, 2019; Aelbrecht & Harwood 2019; Ginzburg & Miller 2019; Poppels 2020; Poppels 2022).

On the basis of contrasts like those exhibited in (1), early generative studies assumed that syntactic identity was necessary for VPE. A mismatched antecedent was claimed to lead to ungrammaticality. (1a) shows the contrast between a case where the antecedent and ellipted VP are both active and a case of voice mismatch, where the antecedent is passive while the ellipted VP is active. Similarly, the contrast in (1b) shows that replacing a syntactically identical VP antecedent by a semantically similar nominal antecedent reduces acceptability.[1]

(1) a. (i) The company fired Jim yesterday. We weren't surprised that they did ~~fire Jim yesterday~~.
(ii) Jim was fired yesterday by the company. #We weren't surprised that they did ~~fire Jim yesterday~~.

b. (i) We were devastated that they attacked our positions. When they did ~~attack our positions~~, we lost all hope.
(ii) We were devastated by their attack against our positions. #When they did ~~attack our positions~~, we lost all hope.

Since the beginning of this discussion, VPE has been contrasted in this respect with Verb Phrase Anaphora (VPA, e.g., *do it*, *do this*, and *do that*), which have been claimed, by contrast, *not* to require syntactic identity with their antecedents as illustrated in (2). As opposed to the previous VPE examples in (1), we see that the voice mismatch in (2a) and the nominal antecedent in (2b) do not seem intuitively to decrease acceptability.

(2) a. Jim was fired yesterday by the company. We weren't surprised that they did it.
b. We were devastated by their attack against our positions. When they did it, we lost all hope.

[1] In order to make the examples easier to read, the antecedent is underlined and the pre-elliptical auxiliary (with VPE) or *do it* (with VPA) is double underlined. In some examples we strike out the putatively ellipted material in the elliptical position. This is a mere expository device intended to clarify interpretation. It is never a suggested syntactic analysis.

In order to explain this difference in behavior, Hankamer & Sag (1976) drew a distinction between **deep** and **surface** anaphors, proposing that VPA were deep anaphors whereas VPE was a surface anaphor. They claimed that deep anaphors simply require a recoverable antecedent in the discourse model, so that they allow both exophoric uses and syntactically mismatched antecedents. Surface anaphors, by contrast, were claimed to require endophoric linking to a syntactically identical antecedent (with the exception of minor inflectional variation), thus excluding mismatched antecendents and exophoric uses.

The deep vs. surface anaphora distinction gave rise to an abundant literature in psycholinguistics, aiming to investigate to what extent experimental evidence could be brought to bear on the topic. Part of this literature concentrates on the specific distinction made between VPE and VPA, seeking to corroborate their status as surface and deep anaphors respectively (see e.g., Murphy 1985, 1990; Tanenhaus & Carlson 1990; Mauner et al. 1995; Roberts et al. 2013; Bélanger 2014; see also Xiang et al. 2019, which compares VPE with Null Complement Anaphora, also claimed to be a deep anaphor by Hankamer and Sag). Other papers specifically focus on the surface status of VPE (e.g., Malt 1985; Arregui et al. 2006; Martin & McElree 2008; Kertz 2010; 2013; Kim et al. 2011; Grant et al. 2012; Kim & Runner 2018; Frazier & Duff 2019; Poppels & Kehler 2018; 2019; see Phillips & Parker 2014 for an overview).

Very early on (starting with Sag's 1976 thesis), it became apparent that the data involved in the deep vs. surface distinction were more complex than initially assumed. Attested examples were found of exophoric uses of VPE and of VPE with syntactically mismatched antecedents, which seemed intuitively quite acceptable and did not appear to be processing errors:

(3) a. This information could have been released by Gorbachov, but he chose not to ~~release this information~~. (Hardt 1993: 37, (131)). (VPE with active VP ellipsed and passive antecedent)

b. [A seeing that B has dyed his hair green] A—You didn't ~~dye your hair green!~~ (Hankamer & Sag 1976: 409, fn.19, (iii)). (Exophoric VPE).

The present paper is centered on VPE with nominal antecedents, which it contrasts with the VPA *do it* in this respect (compare (1b) and (2b) above). The case of nominal antecedents has been less studied (see, e.g., Tanenhaus & Carlson 1990; Kim et al. 2011; Roberts et al. 2013). It remains much less understood than voice mismatches and the literature has provided far fewer attested examples (most of them coming from Hardt 1993: 34–36).

The central dilemma in the debate on whether syntactic identity is necessary in VPE has been the following: if one assumes that it is, then it is difficult to understand why acceptable cases of mismatch, such as those in (3), exist at all; on the other hand, if one assumes that it is not, and that it is only necessary to be able to recover a semantically appropriate antecedent from the

discourse context, then one cannot explain why most constructed cases of mismatches intuitively seem so unacceptable (as in (1a) and (2b) above).[2]

Three central types of solutions have been proposed in the literature to account for the variety of acceptability judgments that arise in the case of VPE with mismatched antecedents. Kehler (2000, 2002) and Kertz (2010, 2013) have proposed that mismatches are always grammatical. However they further suggest that there are general discourse conditions which impact on the use of VPE and that unacceptability results when these conditions are not respected. Kim et al. (2011) suggest that mismatches are grammatical but that they induce additional processing difficulties leading to reduced acceptability (this is made plausible by the large literature showing that parallelism in general—and in particular independently of ellipsis—facilitates processing, see, e.g., Frazier et al. 1984; Frazier et al. 2000; Frazier & Clifton 2001; Dubey et al. 2005.) The latter position is compatible with the first and is in fact accepted by Kertz (2013), who suggests that lack of parallelism can intervene as a second factor reducing acceptability beyond the effect of the violation of discourse constraints. The present paper adopts this combined perspective and shows how it can explain acceptability patterns for VPE with nominal antecedents. A third position, the 'recycling hypothesis', which will be discussed in section 6, has been defended in a series of papers by Frazier and her collaborators, see e.g. Arregui et al. (2006) and Frazier (2013). They claim that VPE requires a syntactically identical antecedent and that any mismatch leads to ungrammaticality. They develop a theory of repairs to such ungrammatical structures that explains why they can be felt to be more or less acceptable in different cases. In a similar vein, Kim & Runner (2018) propose that there are construction-specific grammatical constraints requiring identity but that decreased acceptability can also result from complexity differences in interpretation linked to discourse properties.

In this paper we provide new evidence in favor of a referential theory of VPE (see Poppels 2022), i.e., the idea that VPE does not require a syntactically identitical antecedent, but rather simply requires that an appropriate representation of a referent be recoverable from the discourse model, which can be defined as a mental representation which coherently combines information from the text (linguistic or nonlinguistic utterance acts) and from the nonlinguistic context. In

[2] Some analyses defending syntactic identity make assumptions about syntactic structure such that, for instance, corresponding active and passive VPs are identical at the relevant level of structure (see e.g. Fiengo & May 1994; Johnson 2001; Merchant 2008; Merchant 2013). Such analyses run into exactly the same kinds of problems as those that do not require syntactic identity since, in the absence of further restrictions, they predict that all such cases of mismatch should be acceptable; see Kertz (2013) for discussion. Similarly, some analyses have defended the idea that semantic (rather than syntactic) identity with a linguistic antecedent is the appropriate condition on VPE (see, e.g., Rooth 1992; Merchant 2001; Griffiths 2019a; Stockwell 2020; Stockwell 2023). Such theories also predict that VPE with mismatched antecedents is grammatical (to the extent that the mismatched antecedent is semantically identical to the ellipted VP).

this, we follow the ideas of Webber (1979), Sag & Hankamer (1984) (though they limit this to deep anaphors and do not consider VPE as a deep anaphor), and Cornish (1999).

Thus, the evidence presented here argues in the same direction as other classical evidence that has been adduced against the requirement of syntactic identity on VPE. In particular, the possibility of split antecedents (where the antecedent-trigger does not form a constituent, see, e.g., Webber 1979; Elbourne 2008; Frazier & Duff 2019) and exophoric uses, (see, e.g., Schachter 1977; Miller & Pullum 2014).

The theory of nominal antecedents presented here stems from an example of VPE with a nominal antecedent found in the COCA (the Corpus of Contemporary American English, Davies 2008-):

(4) Mubarak's survival is impossible to predict and, even if he does [~~survive~~], his plan to make his son his heir apparent is now in serious jeopardy. [COCA: CBS Evening News]

This example seemed strikingly natural and raised the question of why it sounded so much better than randomly constructed nominal antecedent cases (as in (1b-ii) above). The current proposal builds on previous work on the construction-specific discourse conditions constraining VPE as opposed to the VPA *do it* (see Miller 2011 and Miller & Pullum 2014). The key insight proposed in those papers is that VPE is more acceptable when there is a salient alternative in the context and the purpose of the utterance containing the ellipsis is to choose one branch of this alternative. What is special about (4) is that the noun *survival* can express a concealed question of a type that has not been previously noticed. Specifically, the first clause in (4) means something close to 'It is impossible to predict whether or not Mubarak will survive', in which the indirect polar interrogative makes the polar alternative 'Mubarak will survive or Mubarak will not survive' salient. The following *even if* conditional clause, containing the VPE, selects one branch of this alternative, as required. Thus the unusual lexical semantics of the noun in this context allows it to satisfy the discourse constraints on VPE in this particular use, making VPE acceptable.

In this paper, section 2 presents the relevant discourse conditions on VPE. We argue that they must be interpreted as preferences and that they are construction-specific (viz., they apply to VPE but not to the VPA *do it*). In section 3, we examine more closely the lexical semantic properties of the relevant class of polar nouns, which can be interpreted as concealed polar questions. Building on these two insights, section 4 shows how they combine to explain the intuitive acceptability of certain cases of VPE with nominal antecedents.

In section 5, we lay out the central experimental data that provide support for our hypotheses. Experiment 1 provides crucial background evidence showing that VPE is preferred to the VPA *do it* when the antecedent is a polar question and the purpose of the subsequent anaphoric clause is simply to provide a yes/no answer (*Yes, she did./No, she didn't.*). The norming study, presented

in section 5.2, corroborates the fact that polar nouns do in fact allow the relevant polar question type readings.

Experiment 2 is the centerpiece of the study. It compares the acceptability of VPE and *do it* depending on two contrasting factors, (i) whether the antecedent is a VP or NP; and (ii) whether the context makes an alternative salient. The results show that when there is an alternative context, VPE is preferred to *do it* both with verbal and nominal antecedents, while the opposite is true in non-alternative contexts. This provides striking corroboration for our proposed discourse conditions and supports the idea that they are construction-specific (since they apply differently to the cases of VPE and *do it*).

However, in alternative contexts, VPE is still judged to be more acceptable with verbal antecedents than with nominal antecedents. Experiment 3 studies the acceptability of the antecedents used in Experiment 2. It shows that, independently of the anaphoric clause, there is an interaction between NP vs. VP and alternative vs. non-alternative contexts. VPs are preferred in alternative contexts while NPs are preferred in non-alternative contexts. We suggest that part of the difference in acceptability between VPE with verbal and nominal antecedents in alternative contexts is in fact due to a spillover effect of the corresponding difference in the acceptability of the antecedent.

Finally, section 6 provides some comparison with Frazier's recycling hypothesis and section 7 concludes, summing up the relevance of our findings for the debate on the role of syntactic identity and discourse conditions in the acceptability of VPE and the VPA *do it*.

2 Discourse conditions on VPE and *do it*

There has been relatively little discussion of the discourse conditions on VPE in the generative literature, see Winkler (2016) and Kehler (2019) for an overview. The first major proposal is that of Rooth (1992) who argues that ellipsis is licenced by redundancy and is akin to overt proforms and prosodic reduction, involving only a difference in the degree of reduction. Kehler (2000; 2002) and Kertz (2010; 2013) follow in this tradition and propose general discourse conditions which can apply to occurrences of VPE. They suggest that the reduced acceptability found in certain mismatches can be attributed to violations of these conditions.[3] Though this line of reasoning has led to significant insights, the specific implementations provided by Rooth and Kertz are in principle incapable of explaining the choice between VPE and VPA such as *do it/this/that*, because, as explicitly stated by Rooth, the general discourse principles involved are expected to apply identically in the case of ellipsis and overt anaphors.[4]

[3] In this context it should be noted that Tanenhaus & Carlson (1990) and Mauner et al. (1995) provide some discussion on the way discourse conditions on passive interact with the acceptability of VPE with voice-mismatched antecedents; similarly, Clifton & Frazier (2010) show that main clause antecedents are preferred for VPE.

[4] Stockwell (2020, 2022), proposes extending Rooth's analysis by adding a requirement for contrast, which he assumes to be specific to VPE (Stockwell 2020: 188, fn.1). However, this idea seems to run into trouble with the very frequent use of VPE to confirm a previous assertion (see example (6c) below).

In fact generative studies opposing VPE and VPA have typically tacitly assumed that there are no differences in their uses beyond those resulting from some version of the deep/surface distinction of Hankamer & Sag (1976) and the fact that *do* in *do it/this/that* is the lexical verb, making stative antecedents impossible (see Culicover & Jackendoff 2005: 283ff and Flambard 2018 for qualifications). Thus, under those assumptions, unless there is no syntactically identical antecedent (in which case only VPA is possible) or the antecedent is stative (in which case only VPE is possible), nothing is predicted at all about the choice between the two.

Miller (2011) and Miller & Pullum (2014) build on this work and propose construction-specific discourse conditions bearing specifically on the felicity of VPE, as opposed to VPA (these are not intended to apply to VPE in comparative structures, which is much less constrained). Following Kertz (2008), they make a distinction between two central uses of VPE, which they call Aux-choice and Subj-choice and propose the following conditions on them (Miller & Pullum 2014: 12):[5]

(5) a. **Type 1: Auxiliary Choice**

FORMAL CHARACTERISTICS: The subject of the antecedent is identical with the subject of the VPE construction and the auxiliary is (at least weakly) stressed, signaling a new choice of tense, aspect, modality, or (in the most overwhelmingly frequent case) polarity.

DISCOURSE REQUIREMENT: A choice between the members of a jointly exhaustive set of alternative situations must be highly salient in the discourse context, and the point of the utterance containing the VPE is strictly limited to selecting one member of that set.

b. **Type 2: Subject Choice**

FORMAL CHARACTERISTICS: The subject of the antecedent is distinct from the subject of the VPE construction, and stressed if it is a pronoun.

DISCOURSE REQUIREMENT: A particular property must be highly salient in the discourse context, and the point of the utterance containing the VPE must be strictly limited to identifying something or someone possessing that property.

These two types of VPE are illustrated in (6) (SMALL CAPITALS indicate stress):[6]

[5] See also Winkler (2000) who proposes a parallel distinction between 'polarity focus' (which she considers the unmarked case of VPE) and 'symmetric focus'. Winkler argues that Rooth's 1992 analysis of VPE only accounts for the latter (Stockwell 2020: 191, fn5, suggests that this stems from an unduly restrictive reading of Rooth). An anonymous referee questions Miller and Pullum's claim that pronouns are different from full NP subjects in specifically requiring stress in Subj-choice cases. This question requires further investigation. Furthermore, it should be noted that there are cases which combine subject and auxiliary choice with a contrastive focus both on the subject and polarity, e.g. *A.—Mark shops in women's. B—Andy doesn't.* These mixed cases are less frequent and their properties remain to be investigated in detail. It appears that in general one of the choices is subordinate to the other, as indicated by greater prosodic prominence on either the subject or the auxiliary.

[6] Kertz (2010, 2013) proposes that the distinction between Aux-choice and Subj-choice is relevant for the acceptability of voice mismatches in VPE: Aux-choice VPE allows active/passive mismatches whereas Subj-choice VPE does not. As will become clear (cf. discussion of example (20) below), the same is true for nominal antecedent mismatches, but for different reasons.

(6) a. A.—Does he shop in women's? B.—He DOESN'T/He DOES. [Aux-choice]
b. A.—Does he shop in women's? B.—#He DOES to find things his size. [Aux-choice]
c. A.—He shops in women's. B.—He DOESN'T/He DOES. [Aux-choice]
d. A.—The guy who shops in women's was here today. B.—#He DOESN'T/#He DOES. [Aux-choice]
e. A.—Mark shops in women's. B.—ANDY does too. [Subj-choice]

In (6a) we see a typical Aux-choice usage: the subject of the antecedent is coreferential with that of the VPE construction, the auxiliary is stressed and signals a choice of polarity. A's polar question sets up the proposition *p*, 'he shops in women's', as the maximal element of QUD (the ordered set of Questions currently Under Discussion, see Roberts (1996), Ginzburg (2012))[7] and makes the alternative $p \vee \neg p$ highly salient.[8] B's answer is limited to choosing one branch of the alternative, conforming to the final clause of (5a). On the other hand, (6b) is intuitively felt to be less felicitous. This is because adding the adjunct *to find things his size* makes B's answer go beyond choosing between the branches of the alternative and resolving the maximal QUD, and thus violates the final clause of (5a).[9] Example (6c) shows that asserting a proposition *p* makes it the maximal QUD (see, e.g., Ginzburg 2012), ensuring that the alternative $p \vee \neg p$ is salient, and allowing either contradiction or confirmation by VPE. In (6d) we see a case where the content of a proposition *p* is not asserted, but introduced as background information that is not under discussion (in this specific case, within a relative clause). As a consequence, the alternative $p \vee \neg p$ is not made salient and *p* is not added to QUD. In such cases contradicting or confirming *p* using VPE is much less felicitous, because it requires major accommodation on the part of the addressee to access the alternative. Finally, (6e) illustrates Subj-choice VPE: the subjects are distinct, the property (or 'open proposition') 'shops in women's' is salient in the discourse context and the purpose of the VPE utterance is to identify someone who possesses this property.[10]

[7] Keshet (2013), Elliott et al. (2014) and Kehler (2015) also defend the relevance of QUD for VPE in the context of what makes sloppy identity readings acceptable; Weir (2014) suggests that QUD is relevant to fragments (i.e., Bare Argument Ellipsis).

[8] Malt (1985) is a precursor on the use of VPE in discourse. Her experimental results can be taken to provide evidence that VPE prefers an antecedent that is the maximal element of QUD.

[9] The acceptability of (6b) is markedly improved if the adjunct is separated by a comma in writing and prosodically detached in speech, because it is then interpreted as an afterthought, and the main content of B's answer is limited to choosing one branch of the alternative. Similarly, as pointed out by an anonymous reviewer, adding a parallel purpose adjunct in the antecedent makes the post-elliptical adjunct contrastive. This in turn makes the ellipsis felicitous: *A.—Does he shop in women's to find his favorite colors? B.—He does to find things his size*. With contrastive adjuncts, this becomes a felicitous case of adjunct-choice pseudogapping (see Miller 2014).

[10] Frazier et al. (2007) provide some support for our condition (5b). They examine cases of what we are calling Subj-choice VPE and give experimental evidence in favor of the idea that, when there are two competing VPs in the discourse context that could both be plausible antecedents, the presence of an L+H* pitch accent on the subject of one of these VPs makes it the preferred candidate for resolving the ellipsis (their experimental materials were constructed so that this was the only pitch accent in the preceding context). Any theory of focus will predict that the contrastive focus on the subject marked by the L+H* pitch accent makes the property expressed by the VP highly salient in the discourse context.

Let us now turn briefly to discourse conditions on the verbal anaphor *do it*.[11] For present purposes, it is enough to know that it is dispreferred in cases where there is a salient alternative and the purpose of the utterance is limited to choosing one branch of the alternative, as shown in (7): *do it* is intuitively dispreferred as a simple answer to a polar question (7a) or as a confirmation or contradiction of a previous assertion (7b). On the other hand, *do it* is preferred if the point of the utterance is not simply to choose one branch of the alternative, as shown in (7c) where a non-contrastive adjunct is present.[12] Its preferences in usage are thus at least in part complementary to those for VPE given in (5).

(7) a. A.—Does he shop in women's? B.—#He DOESN'T DO IT/#He DOES IT.
b. A.—He shops in women's. B.—#He DOESN'T DO IT/#He DOES IT.
c. A.—Does he shop in women's? B.—He DOES IT to find things his size.

To conclude this discussion, we address two issues, the status of the discourse conditions in (5) and the reason for which they must be understood as construction-specific, rather than predictable on general grounds, for instance, from a general theory of focus.

First, it is important clarify that the discourse conditions provided in (5) should not be understood as all or nothing conditions. When it is claimed that a set of alternatives or a property must be 'highly salient' in the discourse context, the notion of salience must be understood as gradable. Higher salience correlates with higher acceptability. The conditions must then be understood as preferences that can be satisfied to a greater or lesser extent (this may not be entirely clear from the formulation in (5) which we quote from Miller & Pullum 2014). As will appear in the results of Experiments 1 and 2 below, the differences in acceptability involved are relatively small, but they are significant.

Second, let us make clear why we believe that the discourse conditions in (5) are construction-specific and cannot be derived from more general principles of discourse, in particular, why they cannot be derived from a general theory of focus. Various studies since Rooth (1992) have attempted to derive conditions on VPE from a theory of focus and accent placement (see, e.g., Stockwell 2022). For instance, in line with Rooth's proposals, one might suggest that condition (5a) would be a consequence of having the nuclear stress on the auxiliary, while (5b) would derive from the nuclear stress on the subject. Tempting though this approach may seem, it is not tenable because it cannot predict the differences in acceptability between VPE and the VPA *do it*. In identical discourse contexts, identical placement of accent across both constructions would be predicted to lead to the same discourse requirements in terms of focus structure. This

[11] For a detailed discussion of the usage conditions on the VPA *do it*, *do this*, and *do that*, see Flambard (2018). Flambard shows that the acceptability of these VPA is affected by various other factors, so that they are not simply in complementary distribution with VPE, as (5) might lead one to expect.

[12] Older corpus-based studies have noticed these preferences: Souesme (1985: 51) notes that *do it* is often modified by an adjunct and Levin (1986: 3) notes that VPE is almost never followed by an adjunct.

makes it impossible to account for differences in acceptability like those between (8a) and (8b) or between (9a) and (9b):

(8) Did she open the door?
  a. She DID.
  b.# She DID it.[13]

(9) Kim opened the door.
  a. PAT did too.
  b.# PAT did it too.

Consider similarly the clause in the discourse requirements of both (5a) and (5b) that claims that VPE must respectively be limited to selecting one member of that set or limited to identifying something or someone possessing that property. By contrast, the VPA *do it* is preferred to VPE in cases where the anaphoric clause goes beyond. These intuitions are corroborated by Experiment 1 below and illustrated by the contrasts in (10).

(10) A.—Did Lily rake up the leaves?
  a. B.—Yes, she DID.
  b. B.—#Yes, she DID it.
  c. B.—#Yes, she did with great CARE.
  d. B.—Yes, she did it with great CARE.

Again, the position of the nuclear accent does not vary according to the choice of VPE or VPA as the anaphoric construction. If there is a non-contrastive adjunct, it will bear the nucleus; if there isn't, the nucleus will be on *did*. Consequently, the focus structure is identical for both VPE and *do it*. This, in turn, entails that general properties of focus cannot explain the choice between the two constructions.

We consequently concur with Stockwell (2020) in claiming that the discourse conditions on VPE must be construction-specific, see footnote 4 above. In fact, it should not be surprising that discourse conditions on syntactic constructions might be construction-specific and interpreted as gradable preferences. Work that has centered on the topic, e.g., Birner & Ward (1998), has presented numerous cases where this is true.

This view of discourse conditions as construction-specific suggests that they can be most naturally formalized in frameworks such as HPSG (e.g., Sag 2012) or Construction Grammar (e.g., Goldberg 2006) which make it easy to state constraints across different levels of representation.[14]

[13] Note that one cannot simply say that *do* cannot be accented in *do it*. In an appropriate context (e.g., where Sally is attempting to jump over a bar), uttering (8b) with exophoric reference to the situation would be perfectly felicitous.

[14] One might consider augmenting these frameworks with a probabilistic component to account for the graded nature of the constraints, perhaps along the lines of Optimality Theory (e.g., Legendre et al. 2001) or Harmonic Grammar (e.g., Pater 2009).

3 A new class of concealed questions

It has been known since Baker (1968) and Grimshaw (1979) that certain NPs, known as 'concealed questions', can have question type meanings when placed in certain interrogative contexts. Classical examples include:

(11) a. Kim knows the price of this book. (≈ *Kim knows what the price of this book is.*)
b. Kim told me the answer to the question. (≈ *Kim told me what the answer to the question is.*)

The crucial properties of concealed questions are (i) they have a question type meaning;[15] (ii) they can occur in most typical indirect question contexts (see Frana 2017: 28–29 for a discussion); and (iii) they have the distribution of noun phrases rather than that of a clausal complement (as demonstrated at length by Grimshaw 1979), specifically, for instance, they can be inverted with an auxiliary (12):

(12) Does [$_{NP}$ your answer to the question] depend on who asks?

Nathan (2006: 21) claims that "insofar as a concealed question denotes a question, that question is an *identity* question, i.e. one of the form *who X is* or *what X is*." A brief review of the literature suggests that this is the current received position (Nathan specifically argues against Baker on this point). However, Jugnet & Miller (2024) show that there is a previously unnoticed class of nouns that can have polar question type meanings in certain interrogative contexts. These are illustrated in the following examples:

(13) a. The outcome of the negotiations largely depends on Anna's participation in the talks. (Compare: …*depends on whether or not Anna participates in the talk*)
b. I am not sure of the committee's approval of the new guidelines. (Compare: *I am not sure whether or not the committee will approve the new guidelines*)
c. It is impossible to predict the witness's cooperation with the police. (Compare: …*to predict whether or not the witness will cooperate with the police*)

Other nouns that allow this kind of use include *acceptance, approval, attack, consent, cooperation, escape, participation, recovery, resignation, surrender, survival, victory, withdrawal*. For convenience, we will call nouns that have this property 'polar nouns'. As illustrated in (13), these nouns can appear in certain indirect question contexts with a meaning similar to that of an indirect polar question.[16] However the range of contexts in which they can appear with such a meaning is far

[15] The precise semantics necessary to account for them is at the heart of most of the recent literature (see Baker 1968; Grimshaw 1979; Nathan 2006; Frana 2010; Barker 2016; Frana 2017; Frana 2020; inter alia).

[16] Note that in (13a) and (13c) the nouns *participation* and *cooperation* can also be understood as classical concealed questions: they allow a constituent question interpretation in terms of manner and/or degree, i.e. they also allow paraphrases of the type 'depends on {the way in which/the degree to which} Anna participates in the talks' (though these are less accessible as shown by the paraphrase judgment norming study discussed below in section 5.2). Nathan

more restricted than for classical concealed questions.[17] It is beyond the scope of this paper to investigate the proper semantic analysis of polar nouns when they are used in these contexts (see Jugnet & Miller 2024 for further information). Given their restricted distribution, it is likely that they do not in fact denote questions. However it is clear that when they are used in the relevant question contexts, they make a polar alternative salient.

It is important to note that the question type interpretation only appears in the relevant subset of contexts allowing indirect interrogatives. In most other contexts polar nouns are interpreted as denoting facts and do not allow *whether*-based paraphrases like those in (13), as illustrated in the following examples:

(14) a. Anna's participation in the talks annoyed us. (≠ *Whether or not Anna participated in the talks annoyed us;* ≈ *The fact that Anna participated* …)
b. The committee's approval of the guidelines enthralled the staff. (≠ *Whether or not the committee approved the guidelines enthralled the staff;* ≈ *The fact that the committee approved* …)
c. John resented the witness's collaboration with the police (≠ *John resented whether or not the witness collaborated with the police;* ≈ *John resented the fact that the witness collaborated* …)

We will suggest below that the factual interpretation of polar nouns is basic, and that the polar interpretation is obtained by coercion, requiring extra processing effort. This is made plausible by the results of Harris et al. (2008), who provide evidence from eye-tracking and MEG studies showing that processing classical concealed questions requires accommodation by the addressee, similar to that found in classical cases of aspectual or complement coercion (see Harris et al. 2008 for references). This is evidenced by the fact that interpreting an NP used as a concealed question requires more processing time and involves higher brain activation than when the same NP is used in a non question environment.

4 VPE with nominal antecedents

We are now in a position to understand why our initial example of VPE with a nominal antecedent, repeated in (15a), is so intuitively acceptable.

(2006) would claim that such interpretations are in fact identity interpretations, i.e. 'the {way in which/degree to which} she participates' could be reduced to 'what her participation is'. It is beyond the scope of this paper to decide whether such a reduction is the best analysis.

[17] If one considers Karttunen (1977)'s typology of indirect question contexts, it seems that, contrary to classical concealed questions, they cannot appear, for instance, after predicates of retaining knowledge (*I don't know the price/#his participation*), predicates of acquiring knowledge (*I found out the price/#his participation*), predicates of communication (*I told them the price/#his participation*). On the other hand, as will appear in the examples provided, they are possible with at least certain predicates of prediction, certain inquisitive predicates, certain predicates of relevance and of dependency.

(15) a. Mubarak's survival is impossible to predict and, even if he does [~~survive~~], his plan to make his son his heir apparent is now in serious jeopardy. [COCA: CBS Evening News]
b. Whether or not Mubarak will survive is impossible to predict and even if he does [~~survive~~], his plan […]

Indeed, *survival* is one of the polar nouns discussed above. It appears in an indirect question type context, as evidenced by the fact that it can be replaced by an indirect polar interrogative, as shown in (15b). In this context, its meaning is coerced into a polar question type meaning, as the paraphrase relation between (15a) and (15b) makes immediately clear. Thus, both the NP *Mubarak's survival* and the clause *Whether or not Mubarak will survive* make the polar alternative {*Mubarak will survive* ∨ *Mubarak will not survive*} salient and the purpose of the elliptical clause is limited to selecting one branch of the alternative, satisfying the discourse condition on Aux-choice VPE given in (5a). Consequently, both the N and the V can serve as antecedents for VPE.

Further constructed examples of the same type are given in (16) to show the robustness of the phenomenon:

(16) a. The integrity of the Senate depends on her participation. If she does, […] (compare: *depends on whether or not she participates*)
b. Since they don't have anyone to replace her with, her resignation is in doubt. If she does, […] (compare: *it is in doubt whether or not she will resign*)
c. The release of this information on the user's part depends on his consent. If he does, […] (compare: *depends on whether or not he consents*)

Furthermore, as shown by the following examples, trying to use *do it* in the same contexts intuitively reduces acceptability, both with nominal and verbal antecedents:[18]

(17) a. The integrity of the Senate depends on {her participation / whether or not she participates}. #If she does it, […]
b. Since they don't have anyone to replace her with, {her resignation is in doubt / it is in doubt whether or not she will resign}. #If she does it, […]
c. The release of this information on the user's part depends on {his consent / whether or not he consents}. #If he does it, […]

This should be compared to the intuitively degraded acceptability of *do it* in answers to polar questions, illustrated in (7a) above.

[18] Replacing *do* by *do it* in (15a) and (15b) intuitively leads to a stronger decrease in acceptability for the independent reason that *survive/survival* do not denote sufficiently actional events to make good antecedents for *do it* in this context (see e.g. Culicover & Jackendoff 2005: 284, Flambard 2018). Note however that adding a means complement (see Souesme 1985: 41) enhances agentivity and makes *do it* more acceptable: *People were intrigued {by the hiker's survival/that the hiker survived}. They couldn't understand how he did it.* In experiments 2 and 3, discussed below, the three items based on *survive/survival* were similarly constructed with a means complement, to avoid independent decrease in acceptability due to lack of agentivity.

It is important to note that most nouns cannot be coerced into polar question type meanings. Some can occur as classical concealed questions, receiving identity question type meanings (or manner/degree type meanings, see footnote 16), as in (18a). For many nouns, no question type interpretation is easily available at all. In all such cases, no polar alternative is made salient and VPE is infelicitous.

(18) a. That depends on her answer. (≈ *That depends on what her answer is*; ≠ *That depends on whether or not she answers*)
#If she does […]
b. When Francis leaves her, his decision is unclear. (≈ *It is unclear what his decision is*; ≠ *It is unclear whether or not he decides*)
#If he does […]

It is also crucial to remember that it is not polar nouns as such which can serve as antecedents to VPE. This is only the case when they are in contexts triggering a salient polar alternative reading, like the indirect interrogative contexts in (16).[19] Otherwise, acceptability is expected to be degraded. Thus, if the antecedent appears in a context where it has its basic factive interpretation, the proposition denoted will be understood to be part of the common ground, and not in QUD. VPE is consequently felt to be infelicitous and *do it* is more acceptable, as illustrated in (19).

(19) a. We were annoyed by her participation in the proceedings. (≈ *by the fact that she participated / by the manner in which she participated*)
We wondered why she #did. [Ok: did it.]
b. His friends were annoyed by his cooperation with the police. (≈ *by the fact that he cooperated / by the manner in which he cooperated*)
They wondered why he #did. [Ok: did it.]

Before closing this section, it is worth mentioning that the discourse condition on Subj-choice VPE (5) is very difficult to satisfy with a nominal antecedent. Consider the following examples:

(20) a. Peter's participation surprised us and #John did ~~participate~~ too.
b. Only/Even Peter's participation surprised us although John did too.

The NP *Peter's participation* in (20a) cannot make salient the open proposition 'x participated'. The variants in (20b) appear to be somewhat more acceptable. The presence of *only* or *even* suggests a choice among a set of individuals and contributes to making the property 'x participated' a bit

[19] The following example from the film *Avatar* provides a different syntactic context in which a polar noun can serve as the antecedent for VPE because an alternative reading ('Will they or won't they cooperate?') is coerced: *I need you to learn about these savages, gain their trust. Find out how I can force their cooperation, or hit 'em hard if they don't.* This is intuitively just as acceptable as the variant with a VP antecedent: *Find out how I can force them to cooperate, or hit 'em hard if they don't.*

more accessible in the context. This makes it easier to accommodate the discourse condition on Subj-choice VPE given in (5a) above, and results in an intuitively somewhat more acceptable sentence.

5 Evidence from acceptability experiments

In order to provide empirical evidence for the hypotheses made in the previous sections, a number of acceptability judgment experiments were run on Prolific and Amazon's mechanical turk using the Ibex platform (Drummond 2014) for online experiments. Experiment 1 provides evidence for our hypothesis that VPE is preferred to *do it* in strict answers to yes/no questions. We then present a norming study verifying that the polar nouns we use in Experiments 2 and 3 actually have a salient polar question type reading in the relevant contexts. Experiment 2 is the centerpiece of this study. It investigates the acceptability of VPE and *do it* for verbal and nominal antecedents in contexts providing a salient alternative or not. In Experiment 3, we present acceptability judgments for the first sentences of the sentence pairs investigated in Experiment 2 in order to see whether verbal and nominal antecedents are judged differently in contexts making an alternative salient or not, independently of the anaphorical or elliptical sentence.

5.1 Experiment 1: Acceptability of VPE and *do it* as answers to polar questions

Recall that an underlying assumption of our hypotheses is that VPE is more acceptable than the VPA *do it* in strictly *yes/no* answers to polar questions. On the other hand adding an adjunct goes beyond a strict *yes/no* answer, which violates the discourse requirement on VPE given in (5a), making *do it* more acceptable than VPE. Experiment 1 was run to compare the acceptability of the VPA *do it* vs. VPE with *do* as answers to polar questions with and without adjuncts. A typical item in its four conditions is presented in (21). **Figure 1** shows a typical stimulus in the experiment. (21)

(21) A—Did Peter dance with you at the party?
B—Yes, he did. [Adj–, Do]
B—Yes, he did it. [Adj–, Do it]
B—Yes, he did to make me happy. [Adj+, Do]
B—Yes, he did it to make me happy. [Adj+, Do it]

5.1.1 Design and methods

20 items were distributed across 4 lists following a Latin Square design, randomly mixed with 24 distractors partly from independent experiments with a similar question-answer structure. Each session started with three practice items to familiarize participants with the task and the structure of the items. 60 participants recruited from Prolific, native speakers of American English, living in the US, judged acceptability (explained in terms of naturalness of the answer in the context

of the question) on a 7 point scale. Participants in this as well as in the following experiments were paid the equivalent of 9 pound sterling per hour. 10 participants with accuracy rates for comprehension questions below or equal .85 were excluded from the analyses.

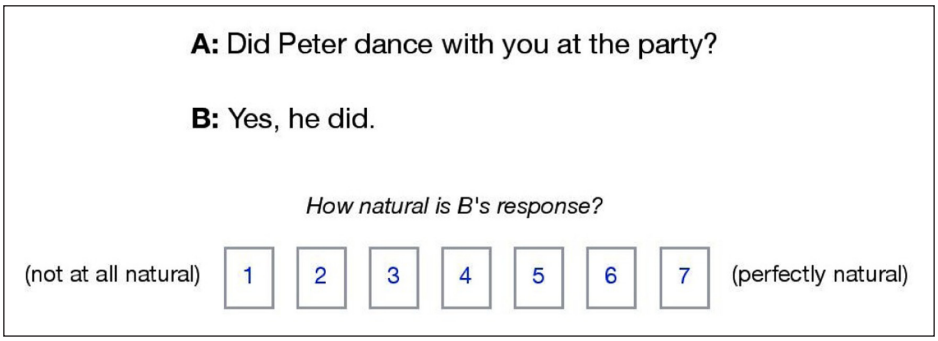

Figure 1: A typical stimulus for Experiment 1 using the Ibex platform.

5.1.2 Results

Condition means for Experiment 1 are given in **Figure 2** (error bars indicate confidence intervals). Model estimates and credibility intervals are given in **Table 1**. We interpret 95% credibility intervals that do not include 0 (0 meaning no effect) as strong evidence for an effect.

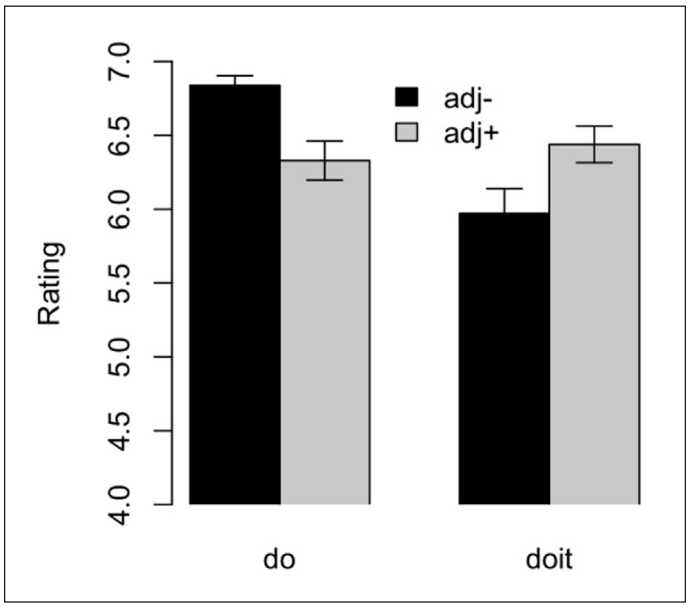

Figure 2: Acceptability of VPE vs. *do it* in polar questions.

	$\hat{\beta}$	95% CrI
Adjunct	–1.11	[–1.99, –0.35]
VPE	–1.87	[–2.82, –0.99]
Adjunct:VPE	4.39	[2.89, 6.12]

Table 1: Model estimates and 95% credibility intervals of model estimates for the fixed factors Adjunct and VPE.

The data were analyzed using R (R Development Core Team 2009). In all experiments reported in this paper, acceptability ratings were submitted to a Bayesian ordinal regression model for inferential statistics (cumulative family) using the bmr function from the R-package bmrs (Bürkner 2017). Mixed effect models were fitted to the data with the fixed factors Adjunct (+,–) and VPE (*do* vs. *do it*), the random factors Participant and Item including random slopes of the fixed effects as well as their interactions. Maximal structure models were used in all experiments (see Barr et al. 2013). Planned comparisons were carried out by fitting a simple model to the data for one level of a condition.

There is strong evidence for an interaction between the two factors ($\hat{\beta} = 4.42$, 95% CrI = [2.95, 6.12], P($\beta > 0$) = 1): the absence of an adjunct increases the acceptability of *do* ($\hat{\beta} =$ –3.34, 95% CrI = [–4.99, –2.01], P($\beta < 0$) = 1) and reduces the acceptability of *do it* ($\hat{\beta} = 1.08$, 95% CrI = [0.33, 1.89], P($\beta > 0$) = .99). Moreover, *do* is generally more acceptable than *do it* (6.61 vs. 5.83, $\hat{\beta} = –1.872617$, 95% CrI = [–2.83, –1.01], P($\beta < 0$) = 1). There is also evidence for slightly lower ratings for sentences with an adjunct ($\hat{\beta} = –1.12$, 95% CrI = [–1.98, –0.34], P($\beta < 0$) = .99).

5.1.3 Discussion

The results corroborate the hypothesis that VPE with *do* is preferred to the VPA *do it* as an answer to a polar question, when the answer is limited to choosing one branch of the alternative. They also provide support in favor of the idea that the acceptability of VPE is reduced when the answer goes beyond simply choosing one branch of the alternative.

We further suggest that the main effects we find, namely (i) the fact that VPE with *do* was more acceptable than *do it* overall and (ii) the fact that the presence of an adjunct reduced acceptability overall, are caused by a property of our materials that we had not considered when we constructed them. Specifically, the answer to the polar question was always introduced by an initial *Yes,* suggesting that the purpose of the utterance was to address the polar alternative raised by the antecedent clause. In this context, given our hypothesis that VPE is better than

do it when it addresses a salient alternative, continuing with VPE is more coherent than than continuing with *do it*, which may explain the overall preference for VPE.[20]

5.2 Norming study: Possible interpretations for polar nouns

A norming study was run to ensure that polar nouns do in fact allow polar question type readings and that they differ in this respect from more typical deverbal nouns. We tested 18 nouns using a metalinguistic task in which participants were asked to judge closeness in meaning between an initial sentence and two or three proposed paraphrases.

5.2.1 Design and methods

Participants saw the initial sentence and the two or three proposed paraphrases at the same time on the screen. Each of the paraphrases was accompanied by a rating scale from 1 to 7.[21] Participants were asked to judge the closeness in meaning of the proposed paraphrases to the initial sentence. Among the paraphrases there was always a polar paraphrase and either a manner paraphrase or an identity paraphrase, depending on which seemed intuitively most plausible. Furthermore, if a degree paraphrase seemed at all plausible, it was also included. (22) provides an example of the test materials with three paraphrases and **Figure 3** shows a typical stimulus with two paraphrases.

(22) a. The outcome depends on Kate's participation. [initial sentence]
b. The outcome depends on how Kate participates. [manner paraphrase]
c. The outcome depends on how intensely Kate participates. [degree paraphrase]
d. The outcome depends on whether Kate participates. [polar paraphrase]

For each of the 18 nouns, three initial sentences were constructed with the nouns respectively as complements of *depends on*, *regardless of*, and *unsure of* (all three of which can govern an indirect interrogative complement clause), as illustrated in (23), giving rise to three conditions of type (22) for each noun and 54 sentences in all. These three versions were distributed across three lists following a latin square design. No effect of the specific sentence context was expected or found empirically.

---

[20] Our hypotheses are further corroborated by a subsequent experiment we conducted, the results of which appear as Experiment 1 in Miller et al. (2020). The conditions of that experiment were the same as those used here, but the materials were different in that they involved subject-choice VPE without the presence of an initial *Yes,* (e.g., *Sue didn't write a song. Sam did (it) (for her)*). The results of that experiment were parallel to those reported here: we found the same strong evidence for an interaction between the two factors, but we found no effect of the choice between VPE and *do it*. We also found an effect of the presence of an adjunct.

[21] We chose not to apply the technique proposed by Ford & Bresnan (2013), who asked participants to distribute points (out of 10) between different constructions, because their technique does not permit one to distinguish interpretations which are all equally close or all equally distinct paraphrases of the original sentence.

(23) a. The outcome depends on Kate's participation.
b. I intend to hold the meeting, regardless of Kate's participation.
c. We are still unsure of Kate's participation.

60 participants judged 18 items, 6 in each condition, among 18 distractors all coming from an independent experiment on attachment ambiguities. The experiment always started with two practice items.

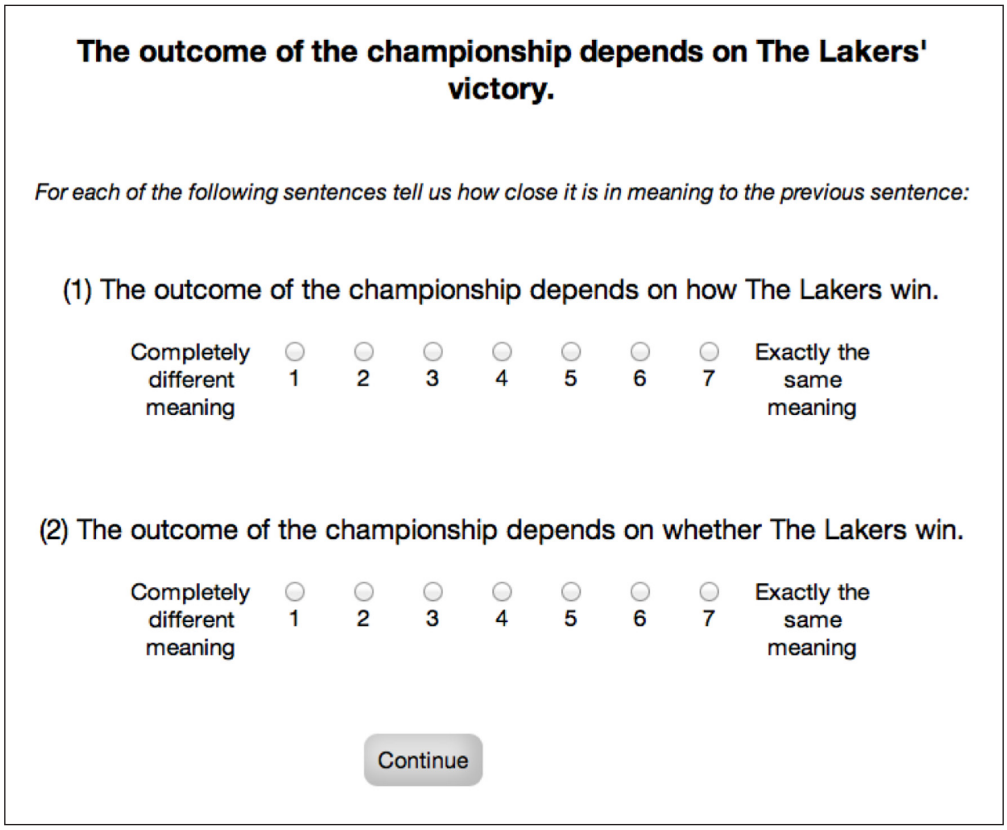

Figure 3: A typical stimulus in the paraphrase study using the Ibex platform.

5.2.2 Results

Table 2 provides the results of the norming study.[22]

[22] Means are followed by the confidence interval in parentheses. The "Intensity" column is empty for those nouns where the intensity paraphrase seemed implausible and was not proposed. For one of the 18 nouns, (*solution*), there was an error in one of the paraphrases that made the results uninterpretable, so that we were forced to eliminate it from the analysis, hence results are provided for the 17 other nouns.

Noun	Manner/Identity	Polar	Intensity
acceptance	3.46(±.61)	6.35(±.39)	
approval	4.31(±.63)	6.60(±.24)	3.73(±.52)
attack	4.27(±.55)	5.58(±.53)	3.73(±.52)
consent	3.71(±.63)	6.46(±.30)	2.92(±.51)
cooperation	4.59(±.57)	6.41(±.32)	3.63(±.54)
escape	2.94(±.47)	5.60(±.51)	
participation	3.71(±.53)	6.60(±.21)	3.56(±.52)
recovery	4.94(±.56)	5.82(±.43)	4.15(±.57)
resignation	2.85(±.51)	6.4(±.30)	
surrender	2.56(±.48)	5.40(±.61)	
survival	3.00(±.53)	6.46(±.29)	
victory	2.52(±.53)	6.6(±.25)	
withdrawal	3.21(±.51)	6.35(±.34)	2.52(±.48)
answer	6.02(±.46)	3.27(±.55)	
decision	6.58(±.31)	3.37(±.59)	
arrival	2.35(±.48)	5.94(±.46)	
destruction	3.33(±.58)	6.29(±.34)	3.81(±.56)

Table 2: Average judgments of quality of paraphrases in the norming study.

Clearly, the thirteen intuitively identified polar nouns (from *acceptance* to *withdrawal*) strongly favor a polar interpretation over a manner/identity interpretation. On the other hand, *answer* and *decision* strongly prefer an identity reading.[23]

This norming study thus leads us to expect that the thirteen nouns that allow polar interpretations should be able to serve as antecedents for acceptable VPE when they are in a context that forces the polar reading. The following subsection discusses the experiments that were run to test this prediction.

5.3 Experiment 2: The choice between VPE and *do it* with nominal and verbal antecedents

The purpose of Experiment 2 was to investigate the acceptability of VPE with *do* and the VPA *do it*, with nominal or verbal antecedents, taking into account whether the context made an alternative salient or not.

[23] It appears that *arrival* and *destruction* allow polar interpretations. However, it was impossible to use these nouns in experiments 2 and 3. *Arrival* raises problems with *do it* because it is insufficiently agentive (cf. fn. 18) and *destruction* was interpreted as passive, leading to VPE with *be* (*We are still unsure of the city's destruction. If it is/*does…*) and making any alternation with *do it* impossible.

5.3.1 Design and methods

Materials were constructed on the basis of three binary factors: (i) Nominal vs. Verbal antecedent (N/V); (ii) Alternative vs. Non Alternative context (Alt+/Alt–); (iii) *Do* vs. *Do it* (Do/Do it). For the nominal antecedents the thirteen polar nouns confirmed by the norming study were used. It should be recalled that the norming study shows that the similarity to a polar paraphrase was judged to be very high for these nouns (the mean is 6.2 on a 7 point scale, where 7 points was explained as 'exactly the same meaning'). This gives us eight conditions per item. A typical item is shown in (24) with its eight conditions.

(24) a. Everyone was annoyed by **Andrew's participation** in the chess tournament. His fans could not understand why he **did**. [Do/N/Alt–]
b. It is impossible to predict **Andrew's participation** in the chess tournament. He is sure to win if he **does**. [Do/N/Alt+]
c. Everyone was annoyed that **Andrew participated** in the chess tournament. His fans could not understand why he **did**. [Do/V/Alt–]
d. It is impossible to predict whether **Andrew will participate** in the chess tournament. He is sure to win if he **does**. [Do/V/Alt+]
e. Everyone was annoyed by **Andrew's participation** in the chess tournament. His fans could not understand why he **did it**. [Do it/N/Alt–]
f. It is impossible to predict **Andrew's participation** in the chess tournament. He is sure to win if he **does it**. [Do it/N/Alt+]
g. Everyone was annoyed that **Andrew participated** in the chess tournament. His fans could not understand why he **did it**. [Do it/V/Alt–]
h. It is impossible to predict whether **Andrew will participate** in the chess tournament. He is sure to win if he **does it**. [Do it/V/Alt+]

For each of the thirteen polar nouns, the Alt+ conditions were constructed using the three contexts given in (25):

(25) a. It is impossible to predict {NP/S[whether]}
b. NP depends on {NP/S[whether]}
c. NP is still unsure {of NP/S[whether]}

The Alt– conditions involved items of the types given in (26):[24]

(26) a. NP was annoyed/embarrased/exasperated/preoccupied/irritated/disgusted/enthralled/amused/fascinated/preoccupied/upset/intrigued/thrilled/dazzled {by NP/S[that]}

[24] Though all the contexts mentioned in (26) are factive contexts, not all factive contexts assign background status to the same degree. A factive predicate such as *was surprised at NP/that S* makes an alternative much more accessible by the very nature of what it is to be surprising: if p is surprising, then $\neg p$ is likely, and stating that p is surprising makes $p \vee \neg p$ salient. An initial version of Experiment 2 had not paid sufficient attention to this, leading to less clear results than those reported here. We thank Anne Jugnet (p.c.) for helping us clarify this issue.

b. NP resented {NP/S[that]}
c. NP was aware {of NP/S[that]}, NP felt threatened {by NP/S[that]}, NP was indifferent {to NP/S[that]}

Forty items[25] were constructed and distributed across eight lists following a Latin Square design and were mixed with fifty-six distractors partly from independent experiments on syntactic and semantic ambiguities. Each experiment started with three practice items. In order to maximize the dispersion of the acceptability judgments, distractors were chosen so as not to contain anything strongly unacceptable. 157 self-identified native speakers judged the acceptability of the second sentence in the context of the first. Cases with reaction times less than 500ms were excluded from the analyses. 0.4% of the observations were affected by this procedure.

Analyses were realized using ordinal regressions from the bmrs package (Bürkner 2017). The model included the predictors Antecedent (noun, verb), Context (Alt+, Alt–) and DoDoit as well as individual random slopes of each predictor as well as their interaction for participants and items. We used uninformative priors, 4 chains and 9000 iterations.

5.3.2 Results

Overall results per condition are given in **Figure 4**. The details of the statistical analysis are provided in **Table 3**.

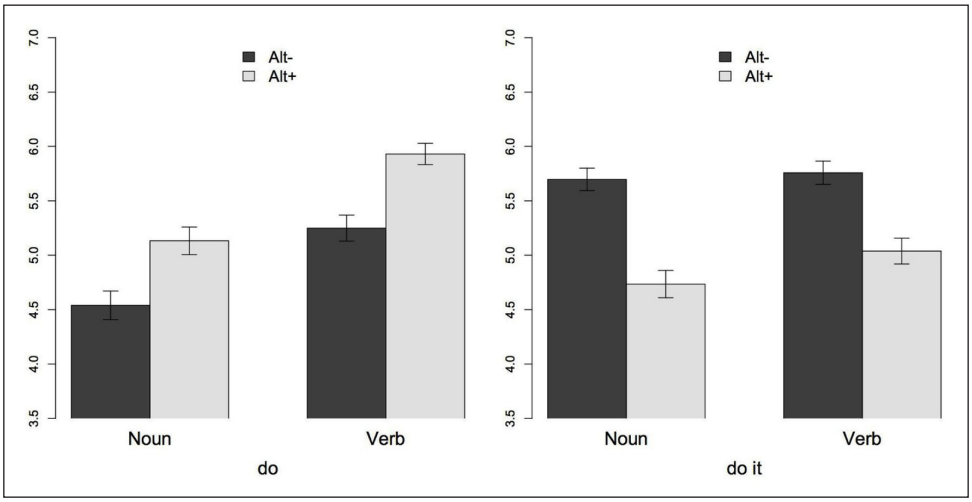

Figure 4: Experiment 2: VPE vs. *do it* with NP vs. VP antecedents in Alt+ vs. Alt– contexts.

[25] As we have 13 polar nouns, using each of these in the three contexts of (25) leads to 39 items. To complete our list of 40, a fourth example using *resignation* in the *depends on* context was added.

	$\hat{\beta}$	95% CrI
DoDoit	0.07	[–0.36, 0.29]
Antecedent	0.71	[0.56, 0.86]
Alt	0.15	[–0.11, 0.41]
DoDoit:Antecedent	–0.88	[–1.15, –0.60]
DoDoit:Alt	2.35	[1.85, 2.84]
Antecedent:Alt	–0.26	[–0.54, 0.01]
DoDoit:Antecedent:Alt	0.15	[–0.27, 0.56]

Table 3: Model estimates and credibility intervals for the fixed factors DoDoit, Antecedent Type (Noun, Verb) and Context(Alt+/–).

Among the major results one can note strong corroboration of the hypothesis according to which VPE with *do* is more acceptable when there is a salient alternative in the context while *do it* is more acceptable when the antecedent is not under discussion. This is shown in **Figure 5** and evidenced by the high probability of an interaction of Do/Doit and Alt+/– ($P(\beta < 0) = 1.00$). VPE was most acceptable in Alt+ contexts ($\hat{\beta} = -1.11$, 95% CrI = [–1.41, –0.81], $P(\beta < 0) = 1$) while it is less acceptable for Alt– contexts ($\hat{\beta} = 1.23$, 95% CrI = [0.88, 1.59], $P(\beta > 0) = 1$).

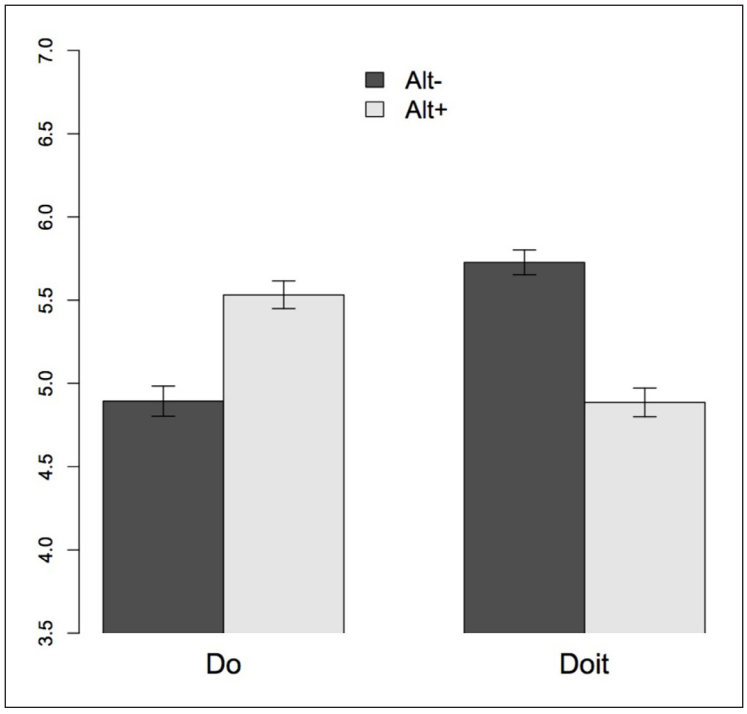

Figure 5: Experiment 2: VPE vs. *do it* in Alt+ vs. Alt– contexts.

We can also confirm, as expected, that *do* in VPE is judged more acceptable with a verbal antecedent than with a nominal antecedent ($\hat{\beta} = 1.10$, 95% CrI = [0.86, 1.34], $P(\beta > 0) = 1$). But as shown in **Figure 6**, this seems to be true for *do it* too, though to a numerically lesser extent ($\hat{\beta} = 0.26$, 95% CrI = [0.12, 0.42], $P(\beta > 0) < 0.99$; but see Discussion of Experiment 3). Those differences lead to a reliable interaction (Do/Doit by Antecedent: $P(\beta < 0) = 1.00$).[26]

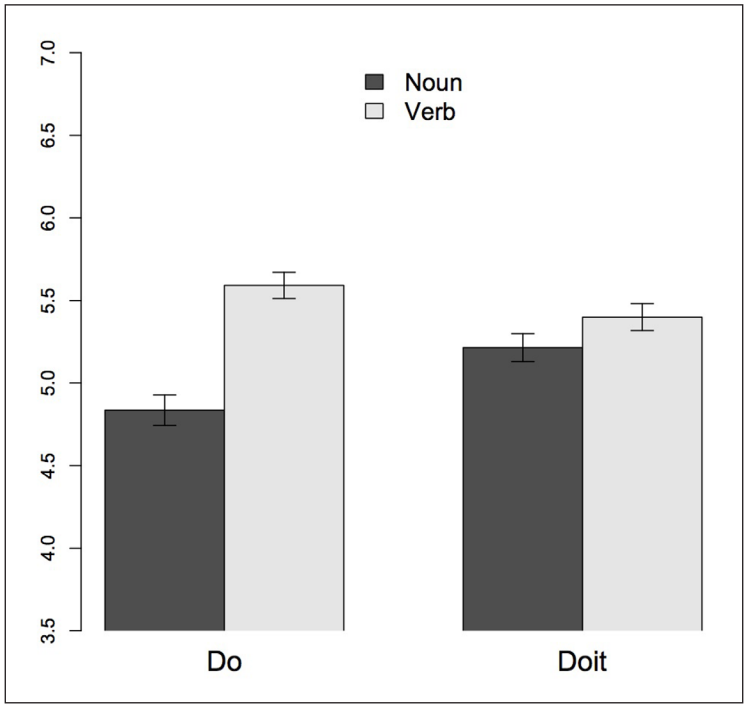

Figure 6: Experiment 2: VPE vs. *do it* and V vs. N.

Looking more closely at the average acceptability of the various conditions of Experiment 2, as shown in **Figure 4**, it is worth making the following points. First, the condition "Do/N/Alt+" is rated at 5.15 (out of 7), a surprisingly high rating in view of the N vs. V mismatch

[26] These results should be compared to those of Tanenhaus & Carlson (1990)'s experiment 2, in which they use a 'make sense' task to investigate differences between VPE and *do it* with verbal and nominal antecedents. They found that reaction times were similarly increased in the mismatch cases, both for VPE and *do it*, which conforms well with the reduced acceptability of nominal antecedents in both cases evidenced here in **Figure 5**. It should be noted that they found a significant difference in the results for the make-sense task, in the case of mismatch, between *do it* and VPE (respectively judged to be 'sensible' in 86% and 71% of cases). We would interpret this result as due to the fact that the discourse conditions on VPE are not satisfied in their nominal antecedent cases. Consider for instance their example (9): "The mention of her sister's name always annoys Sally. However, Tom did (it) anyway out of spite." Note that two factors contribute to making VPE far less acceptable than *do it*, independently of the category mismatch: (i) the event of mentioning is backgrounded by the nominalization and (ii) there is an adjunct after the anaphor.

involved. To put this into context, the average acceptability rating for the whole experiment (which, as mentioned above, involved no strongly unacceptable sentences), including fillers, was 4.73. This should be compared to the condition "Do/V/Alt+", which was rated most acceptable at 5.94. Also of interest is the fact that compared to the condition "Do/N/Alt+" (5.15), there is weak evidence that the condition "Do it/N/Alt+" was less acceptable at 4.77 ($\hat{\beta} = 0.07$, 95% CrI = [–0.10, 0.246], P($\beta >$) = 0.79). This goes against the classical expectation that *do it* (a deep anaphor) should be systematically better than VPE (a surface anaphor) with a nominal antecedent. The condition "Do it/V/Alt+" was rated numerically lower than condition "Do/N/Alt+" at 5.07, but we found strong evidence that it is more acceptable than the condition "Do it/N/Alt+" ($\hat{\beta} = 0.59$, 95% CrI = [0.46, 0.72], P($\beta > 0$) = 1). In sum these data discredit the idea that *do it* should systematically be judged better with nominal antecedents than VPE.[27]

5.3.3 Discussion

We thus believe that it becomes very difficult to claim that the condition "Do/N/Alt+" is ungrammatical (we will come back to this in section 6). However it would be interesting to have some insight into why, with VP Ellipsis in alternative contexts, the nominal antecedent cases are judged less acceptable than the verbal antecedent cases. What can explain this difference in acceptability if it is not due to a difference in grammaticality? A partial answer to this question might be suggested by the results of Harris et al. (2008), mentioned above, which showed that NPs require longer processing times when they are interpreted as classical concealed questions, due to the need for accommodation. This suggests that polar nouns might require accommodation when they express a polar question and consequently be more difficult to process, which might affect acceptability judgments on sentences involving them. Acceptability and processing cost are not necessarily inversely correlated since acceptability can also be affected by constraint violations independent of processing cost. Hofmeister et al. (2014) show evidence that not all acceptability judgments are affected by measures of cognitive capacity such as reading span measures which are typically related to processing cost. However, while acceptability judgments can be affected by other factors than processing cost, increased

[27] Though this is only anecdotal evidence, it is also worth pointing out that for two out of forty items the condition "Do/N/Alt+" was actually judged better than the condition "Do/V/Alt+", and for three items the condition "Do/N/Alt+" was judged only .1 or less worse than the condition "Do/V/Alt+". The following example provides the data and ratings from one of the items which was judged on average (by 20 subjects) to be more acceptable with the nominal variant of the antecedent.

i. We have talked about a possible adoption, but we are still unsure of her consent. [Do/N/Alt+]
We have talked about a possible adoption, but we are still unsure whether she will consent. [Do/V/Alt+]

ii. It will be better for the baby if she does. N:5.74; V:5.26

processing cost typically (though not always) leads to lower acceptability (Hofmeister et al. 2013). The data on processing cost reported in Harris et al. (2008) would thus predict lower acceptability judgments for NPs that have to be interpreted as concealed questions. In order to test this hypothesis, we ran Experiment 3.

5.4 Experiment 3: Antecedent acceptability independent of ellipsis

5.4.1 Design and methods

Experiment 3 used as materials the antecedents of Experiment 2 *without* the following anaphoric or elliptical sentence (the stimuli of Experiment 2 were intentionally constructed to make this possible). This leaves us with two binary factors, N/V and Alt+/Alt–. The task for participants and general procedures were identical to Experiment 2. The forty experimental items were distributed across four lists following a Latin Square design and mixed with fifty-two distractors partly from independent experiments on syntactic and semantic ambiguities. Each experiment started with three practice items. A typical item in its four conditions is given in (27):

(27) a. Everyone was annoyed by **Andrew's participation** in the chess tournament. [N/Alt–]
b. It is impossible to predict **Andrew's participation** in the chess tournament. [N/Alt+]
c. Everyone was annoyed that **Andrew participated** in the chess tournament. [V/Alt–]
d. It is impossible to predict whether **Andrew will participate** in the chess tournament. [V/Alt+]

10 out of 80 participants, recruited on Amazon Mechanical Turk, were excluded as self-reported non-native speakers. Observations with reaction times lower than 500ms were excluded from the analyses. 1% of the observations were affected by this operation.

5.4.2 Results

Antecedent type (Noun vs. Verb) and Alternative (Alt+ vs. Alt–) were included as centred predictors as well as the interaction between the two variables. In addition to random intercepts of items and participants, random slopes for Antecedent and Alternative, and their interaction were included. We used uninformative priors, 4 chains and 9000 iterations. Overall results per condition are given in **Figure 7**. **Table 4** shows estimates as well as credibility intervals for the fixed factors.

We found strong evidence for an interaction effect between antecedent type and alternative ($P(\beta) < 0 = 1$). Verbal antecedents were more acceptable than nominal antecedents in Alt+ ($\hat{\beta} = 0.95$, 95% CrI = [0.51, 1.40], $P(\beta >) = 0.99$). By contrast, nominal antecedents were more acceptable in Alt– ($\hat{\beta} = –0.50$, 95% CrI = [–0.82, –0.180], $P(\beta <) = 0.99$).

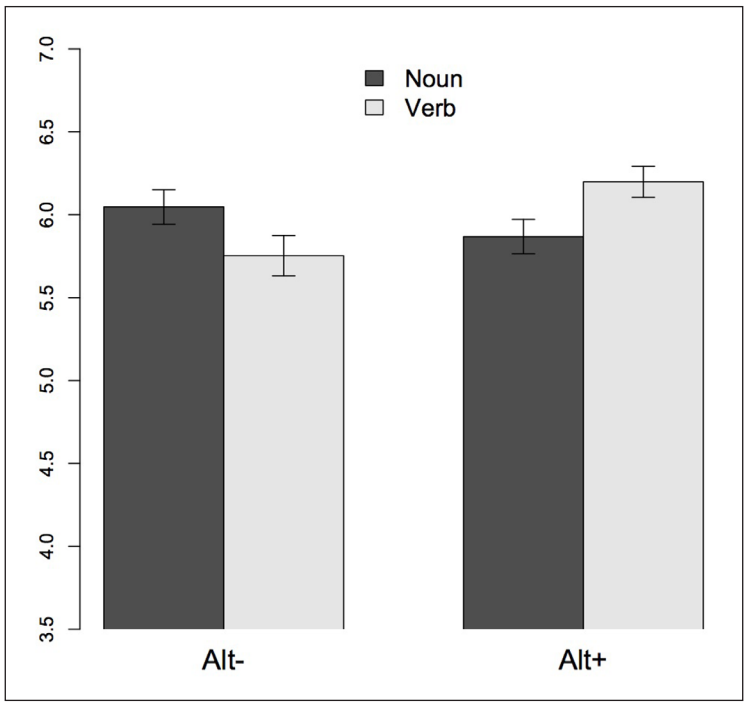

Figure 7: Experiment 3: Antecedents.

	$\hat{\beta}$	95%*CrI*
Antecedent	0.18	[–0.08, 0.44]
Alt	–0.21	[–0.57, 0.14]
Antecedent:Alt	–1.42	[–1.96, –0.90]

Table 4: Model estimates and 95% Credibility Intervals for Antecedent Type (Noun, Verb) and Context (Alt+/–).

5.4.3 Discussion

The decreased acceptability of sentences with polar nouns in alternative contexts suggests that they may indeed be harder to process when they have a question type meaning than when they have a fact type meaning, supporting the idea that some form of coercion might be required. It is interesting to note that there is a symmetric effect for the Alt– environment, where verbs are judged less acceptable than nouns.

In order to find out how far independent processing difficulty of the context sentences in Experiment 2 may have contributed to the effects we found in acceptability judgments in that experiment, we established 'ellipsis-only' ratings by calculating the residuals of the

linear regression between the ratings of the antecedents and ratings of the ellipsis in each context using a linear mixed effects model with random intercepts for participants and items. Although we explicitly asked participants to judge the target sentence in the context of the preceding sentence, we cannot exclude spillover effects from the context sentence itself affecting the judgment of the target sentence. The availability of a context can influence acceptability judgments even if only judgments of the target sentences are elicited (Bernardy et al. 2018).[28]

Figure 8 shows the residual acceptability judgments and **Table 5** gives the fixed effects of the linear regression model (family = gaussian) that we fitted to these data.[29] As before, we ran a maximal Bayesian model such that the models were parallel to those of Experiment 2. Similar to the uncorrected ratings, we found a strong effect of Antecedent for the residuals with verbal antecedents generally leading to slightly higher ratings ($P(\beta > 0) = 1$). We also established strong evidence for interactions of Do/Do it and Antecedent on the one hand ($P(\beta < 0) = 1$) and Do/Do it and Alternative on the other hand ($P(\beta > 0) = 1$). For VPE, Alt+ contexts are more acceptable than Alt– contexts ($\hat{\beta} = 0.836$, 95% CrI = [0.579, 1.085], $P(\beta > 0) = 1$), for Do it, on the other hand, Alt– contexts are more acceptable than Alt+ contexts ($\hat{\beta} = 0.836$, 95% CrI = [0.579, 1.085], $P(\beta > 0) = 1$). Nominal antecedents are moreover less acceptable than verbal antecedents for VPE ($\hat{\beta} = 0.181$, 95% CrI = [0.08, 0.283], $P(\beta >) < 0.99$) while they are more acceptable for Do it ($\hat{\beta} = 0.181$, 95% CrI = [0.08, 0.283], $P(\beta >) > 0.99$). Importantly, even after taking into account the acceptability of antecedents, there is still strong evidence for a difference between Noun and Verb Antecedents for VPE in Alt+ contexts ($\hat{\beta} = 0.74$, 95% CrI = [0.55, 0.93], $P(\beta > 0) = 1$).

In conclusion, Experiment 3 shows that nominal antecedents, in contexts where they have their polar question interpretation, are less acceptable than the corresponding verbal antecedents. Since this effect was established independently of ellipsis, it may contribute to the difference in acceptability of VPE found in Experiment 2 between nominal and verbal antecedents in alterative contexts, as a spillover effect. However, the residuals analysis still showed a disadvantage for nominal antecedents with VPE, which can thus not be explained by a potential context effect.

---

[28] Since there are a variety of possible ways to calculate residuals which may each under- or overestimate the role of the context, the following line of argumentation can only provide a first indication of how to disentangle effects of context and target sentences. Acceptability judgments only provide a single measure for all components of complexity in the two-sentence texts in our experiments. Antecedent ratings significantly predicted judgments in our model: $\hat{\beta} = 0.1315$, se = 0.049, $p < 0.01$.

[29] Note that effects for the residuals will be negative when the judgments are lower than can be expected from the first sentence alone and positive when the judgments are higher than what can be expected from the first sentence alone.

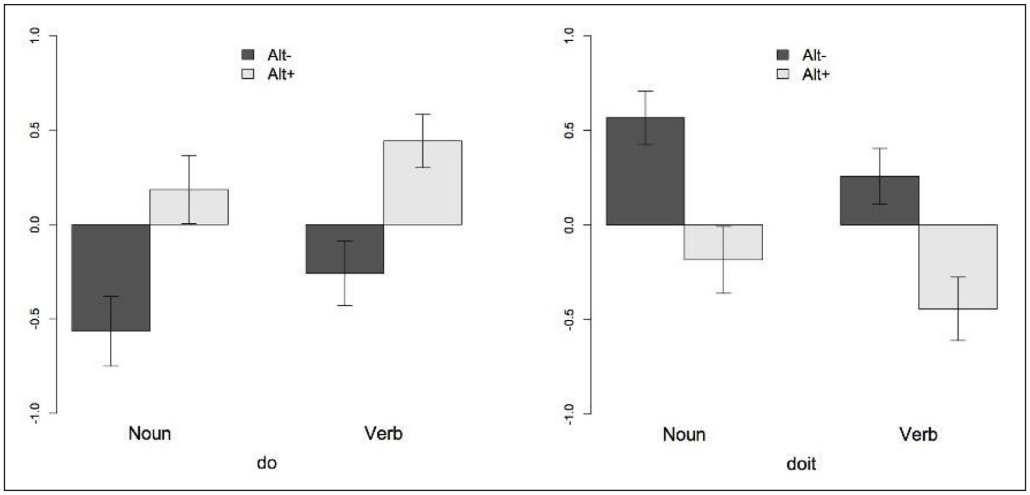

Figure 8: Experiment 3b: Residual acceptability judgments.

	$\hat{\beta}$	95%*CrI*
DoDoit	0.093	[–0.051, 0.236]
Antecedent	0.462	[0.368, 0.560]
Alt	0.113	[–0.072, 0.297]
Antecedent:Alt	–0.065	[–0.251, 0.119]
DoDoit:Antecedent	–0.563	[–0.746, –0.380]
DoDoit:Alt	1.446	[1.143, 1.757]
DoDoit:Antecedent:Alt	–0.151	[–0.437, 0.139]

Table 5: Model estimates and credibility intervals for the fixed factors DoDoit, Antecedent Type (Noun, Verb) and Context(Alt + /–).

6 Comparison with the recycling approach to mismatch in VPE

As has been previously stated, we adopt a position very similar to that of Kim et al. (2011) and Kertz (2013), rather than the recycling theory of Frazier and her collaborators (e.g., Arregui et al. 2006; Grant et al. 2012; Frazier 2013). The latter propose that VPE requires a syntactically identical antecedent and that any mismatches lead to ungrammaticality. However, in some cases an ungrammatical string can be judged to be acceptable (or at least marginally acceptable) because the addressee can use repair strategies to build an appropriately identical antecedent from the actual, mismatched antecedent. A central assumption they make is that the degree of unacceptability is a function of (i) the number of operations required to repair: the more repair operations are needed, the more the example will be judged unacceptable; and (ii) the amount of

evidence available for the repair. In this light, a proponent of recycling might adopt our discourse conditions on VPE and argue that they are compatible with a repair approach. For instance, it could be assumed that it is easier to repair a polar noun like *participation* in Alt+ contexts, where the discourse conditions are satisfied, as opposed to the Alt– contexts, predicting that the former are more acceptable than the latter. This would lead to an alternative recycling-based account for the patterns of acceptability found in Experiment 2.

However, it seems to us that the plausiblity of the recycling hypothesis, as a solution to the variable acceptability of VPE with mismatched antecedents, is crucially linked to the relatively low acceptability ratings reported by Arregui et al. (2006) for nominal antecedents. Specifically, in their Experiment 3, they request a binary judgment of acceptable or unacceptable and they find that VPE was judged acceptable 35% of the time with nominal gerunds[30] as antecedents, as in (28a), and 58% of the time with verbal gerund as antecedents, as in (28b). It is of course difficult to compare the results of their very different experimental protocols (binary judgments) and materials (verbal and nominal gerunds) with those reported on here. However, in the light of the preceding discussion, it is plausible to attribute these much lower acceptability results to the fact that their stimuli do not respect the condition on Aux-choice VPE.[31] Consider, for instance, the specific examples they use to illustrate their materials:

(28) a. Tomorrow night's slow singing of the arias will be difficult/but Maria will. [Nominal gerund]
b. Singing the arias slowly tomorrow night will be difficult/but Maria will. [Verbal gerund] (Arregui et al. 2006: 238, their (14b,d))

The question of whether or not the arias will be sung tomorrow night is not at stake in the antecedent and no alternative is made salient. The most salient QUD is the difficulty of singing the arias; the fact that they will be sung is backgrounded and is not under discussion. Thus acceptability can be expected to be low, regardless of the category of the antecedent and a further penalty for the mismatched nominal gerund is to be expected.

In this light, it seems to us that recycling theory becomes less plausible as an analysis for mismatched antecedents in VPE on general grounds of theoretical parsimony ('Ockham's razor'). In our Experiment 2, mismatched antecedents in favorable discourse conditions are judged to be more acceptable than the average of all sentences (including fillers, none of which were strongly

[30] There is much terminological variation for these forms. In Huddleston et al. (2002): 80ff, for instance, verbal gerunds are called 'gerunds' and nominal gerunds are called 'gerundial nouns'.

[31] This probably goes a considerable way in explaining that verbal gerunds were only judged acceptable 58% of the time. When the discourse conditions on VPE are met, verbal gerunds appear to be completely acceptable antecedents, e.g. *The success of tonight's performance depends on her singing the aria slowly. If she does, there won't be any problems* (this should be contrasted with their example cited here as (28a)). This is not surprising given that the verbal gerund combines with its complements and adjuncts to form a VP, see, e.g., Pullum (1991).

unacceptable). Since putatively repaired sentences are judged as acceptable as grammatical sentences, there is no longer any reliable means to distinguish the two statuses and it is hard to see what might be gained by saying that some equally acceptable sentences are grammatical while others are ungrammatical but repaired.

More importantly, our data raise a crucial specific difficulty for the application of the recycling hypothesis to matched antecedents in VPE. If we accept that encountering a case of VPE first triggers a search for a syntactically identical antecedent and only takes into account various discourse-based strategies in the repair process, there is no reason to assume that the discourse requirements should exert any effect on verbal antecedents. However, in our experiments, contexts with salient alternatives increased acceptability for verbal antecedents as much as they did for nominal antecedents. This effect shows that the discourse conditions affect acceptability independently of repair processes.

More generally, a systematic dispreference for certain antecedents (e.g. non-topics, clefted antecedents) in within sentence pronoun resolution (see Colonna et al. 2012; De la Fuente & Hemforth 2013), or negated antecedents within or across sentences (Shuval & Hemforth 2008) has usually been taken as evidence for reduced accessibility or prominence of the antecedent but never as evidence for ungrammaticality, even though preferences depend on language specific options in the grammar (Hemforth et al. 2010). We propose that the slightly reduced acceptability of nominal antecedents (when opposed to verbal antecedents in a favorable context) should be seen in a similar way.

In a different line of work within the recycling hypothesis framework, Grant et al. (2012) propose that what they call 'Non-Actuality Implicatures' (NAIs), conveyed by non-epistemic modals like *should*, facilitate processing of ellipsis in general, and specifically in cases of voice mismatches. They suggest that the improvement linked to the presence of an NAI is due to the fact that NAIs introduce an alternative (between a desired and actual state of affairs). This triggers a potential QUD and ellipsis is claimed to be easier when the ellipsis clause comments on the QUD. As Grant et al. do not study nominal antecedents, but rather voice mismatches, the details of their results are not directly comparable to ours.

In fact, the NAI hypothesis seems to be a special case of the more general condition proposed by Kertz (2013), namely that it is contrastive topic ellipses that are sensitive to argument structure mismatches such as voice mismatch, while ellipses with Aux-choice (of which NAI cases form a subtype) allow them. In any case, both Kertz's and Arregui et al.'s proposals are in principle incapable of predicting the differences in behavior we have found between VPE and the VPA *do it*. This is explicitly stated by Kertz (cf. the quote provided at the end of the first paragraph of section 2). It is implicit in the discussion of Grant et al., who claim that "the NAI hypothesis predicts that mismatch ellipsis will be more acceptable when there is an NAI, because the alternatives introduced by the NAI will implicitly focus the antecedent making

the antecedent easier to identify and therefore repair" (Grant et al. 2012: 328). Both analyses then attribute the increase in acceptability linked to an NAI to a general principle of anaphor resolution (anaphors are easier to resolve if their antecedents are easier to identify). This then predicts that NAIs should facilitate the resolution of the VPA *do it* with mismatched antecedents just as much as for VPE. Thus, neither Kertz nor Grant et al. can explain the differences found in Experiment 2 between VPE and VPA.[32]

7 Conclusion: VPE and identity

In this section, we sum up the import of the experimental results presented above for the general question of whether syntactic identity is required for VPE and we discuss discuss the broader theoretical consequences of our findings.

Recall that the idea that VPE requires a syntactically identical antecedent (see e.g., Hankamer & Sag 1976) was introduced into the literature to account for the fact that cases with mismatched antecedents are typically less acceptable than those with matched antecedents. In this context, the discovery of intuitively acceptable naturally occurring cases of VPE with mismatched antecedents, in particular due to Hardt (1993), raised a problem. Proponents of syntactic identity theories typically responded to this difficulty by resorting to more abstract syntactic representations, in which, for instance, a nominalization contains a VP, making it possible to satisfy syntactic identity with a nominal antecedent (e.g. Johnson 2001; van Craenenbroeck 2017, who adapt the analysis proposed by Fu et al. 2001 for *do so* with nominal antecedents).

The problem for such theories is that, as such, they predict that it should generally be the case that there is no effect of surface mismatch on acceptability as long as the abstract syntax conforms to the identity constraint. But this leads them to overlook the relevant point at the source of the identity constraint, namely that violations of identity typically reduce acceptability. Thus, the core problem of the variability in the acceptability of VPE with mismatched antecedents is not addressed.

As we discussed in section 1, three types of strategies have been proposed to account for variable acceptability under mismatch: (i) general discourse conditions (e.g., Kehler 2002; Kertz 2013); (ii) processing difficulties (e.g., Kim et al. 2011); (iii) and recycling, which we argued against as a solution to the mismatch problem in the previous section.

---

[32] Before closing this section, we should also mention the eye-tracking study of Roberts et al. (2013) who argue that VPE requires more computation at the ellipsis site than *do it* in the case of voice mismatches and nominal antecedents. However, their results are difficult to interpret because their stimuli typically do not satisfy the discourse conditions on VPE. Consider for instance, their (12c): *The robbery at Citibank was an act of desperation. The police haven't figured out who did*. Here, the first sentence presents the fact that Citibank was robbed as backgrounded information, and not a QUD. Thus the greater processing times they find may well simply be due to the fact that no antecedent satifying the discourse conditions on VPE is accessible.

The present paper is original in that it suggests that beyond the relevance of general discourse contraints, it is also necessary to take into account the construction-specific discourse constraints on VPE proposed by Miller & Pullum (2014). Our Experiments 1 and 2 support their proposals for Aux-choice VPE, showing that it behaves differently from the VPA *do it*, in ways that do not appear to be explainable in terms of the general discourse constraints proposed in the literature.

Specifically, Experiment 1 provides support for the idea that a clause with Aux-choice VPE is more acceptable when it does not go beyond addressing the most salient alternative provided by the discourse context. Experiment 2 corroborates the idea that a clause with Aux-choice VPE is more acceptable when it addresses a salient alternative in the discourse context.

Moreover, with Experiment 2, we established that, when the discourse conditions on VPE are satisfied, VPE with nominal antecedents can be surprisingly acceptable, in fact significantly better than cases of *do it* with a nominal antecedent for which the discourse conditions are not satified (standardly assumed to be fully grammatical) and only slightly less acceptable than the corresponding cases with a verbal antecedent. Further, Experiment 3 showed that subjects rated the acceptability of the nominal antecedents in the materials of Experiment 2 as less acceptable than their verbal counterparts when the antecedent clauses were judged independently. We suggested that part of the small decrease in acceptability between verbal and nominal antecedents in Experiment 2 might in fact be due to the difficulties subjects have in judging the VPE clause independently of the intial antecedent clause.

To sum up, we argue that VPE simply requires accessing an antecedent that satisies the relevant discourse constraints, independently of its category. We assume that the heuristics for establishing a possible candidate antecedent makes use of all the information available in the discourse model as well as available syntactic structure. Syntactically identical antecedents are simply easier to find and are consequently typically more acceptable.

In this perspective, one might speculate about the reasons for which syntactic identity is so much more frequent for VPE than for *do it* in actual usage. We suggest that this results from a combination of two factors. First, syntactically identical verbal antecedents are much more frequent because nominal antecedents only rarely fulfill the discourse requirements on VPE (the same is true for argument structure mismatches as shown by Kertz 2013). Second, VPE requires a more accessible (in the sense of Ariel 1990; 'in focus' in the terminology of Gundel et al. 1993) antecedent than *do it*, as shown by Miller (2011). As a consequence, when there is an identical antecedent, it is much more likely to still be active in short term memory in the case of VPE than in the case of *do it*. Taken together, these two factors entail that searching for a syntactically identical antecedent is a more dependable heuristic for VPE than for *do it* because it leads to success more often.

We speculate that this is what has led VPE to be identified as a surface anaphor, requiring syntactic identity, as opposed to *do it*. If the heuristic strategies involved in recovering antecedents

give more weight to syntactic structure in the case of VPE than in the case of *do it*, all other thing being equal, the absence of a syntactically identical antecedent will reduce acceptability more strongly for VPE than for *do it*. This idea is corroborated by the fact that *do so* also requires a highly accessible/in focus antecedent (see Miller 2011, 2013) and that it has also been identified as a surface anaphor, requiring syntactic identity, despite sharing the presence of main verb *do* with *do it*, contrary to the auxiliary *do* of VPE.

To conclude, the general idea that VPE requires a syntactically identical antecedent has been made much less plausible by recent work on split antecedents and exophoric VPE. For split antecedents (see Webber 1979; Hardt 1999 who use them as arguments against the identity condition and Elbourne 2008 for an opposing view), Frazier & Duff (2019) report a paraphrase elicitation experiment (their experiment 1) showing that, when presented with a sentence like *Wendy is eager to sail around the world and Bruce is eager to climb Mt. Kilimanjaro, but neither of them has so far* and asked to complete a prompt of the type *So far, neither of them …*, subjects typically (69% of responses) produced a superordinate category completion (e.g., *has gone on their anticipated adventures*), inferred from the antecedent, but not in any way syntactically identical to it.

For exophoric uses, Miller & Pullum (2014) show that they are much more common in corpora than usually assumed and that they are perfectly natural as long as the discourse conditions on VPE are satified.

In the same vein, Geiger & Xiang (2021) show that, with VPE, the addressee can access information from the non-linguistic context to specify an antecedent beyond what is linguistically given, even when what *is* linguistically given forms an adequate antecedent on its own. As they point out, this cannot be explained if accessing non-linguistic context is an option only used as a last resort to repair non-identical antecedents or missing antecedents (as suggested by Merchant 2004). The presence of a plausible linguistically given antecedent should preclude any access to the non-linguistic context.

Following up on these results, the present paper provides experimental support corroborating the discourse conditions on VPE proposed by Miller & Pullum (2014) and shows how they allow an account of the variable acceptability of a range of cases of VPE with nominal antecedents. Again, it is unclear how any theory of VPE based on an identity constraint could account for these data.[33]

[33] On the other hand, the data presented in this paper do not impinge on the question of whether or not there is unpronounced syntactic structure in the ellipsis site. Merchant (2001; 2013) and many others have argued in favor of this idea while, e.g., Jacobson (2016) has argued against it. Though syntactic identity and unpronounced structure are often taken to go hand in hand, there is no logically necessary connection between them (for instance, Merchant (2001) argues for unpronounced syntactic structure but imposes a semantic, rather than syntactic, identity condition).

The discourse conditions proposed here hinge on the relation between the elliptical clause and the QUD established by the preceding antecedent utterance. In this way, they enter into a recent tradition that has been investigating the link between ellipsis and QUD. This can be seen as stemming from the proposals of Rooth (1992) and includes such references as Keshet (2013); AnderBois (2014); Elliott et al. (2014); Weir (2014); Kehler (2015); Griffiths (2019b).

Finally, one might wonder why attested examples of VPE with nominal antecedents are rare. Might this not be a strong argument in favor of syntactic identity? We claim that this papers shows exactly the opposite. The problem with nominal antecedents is not a problem of syntax, but rather a problem of discourse constraints. VPE with nominal antecedents is rare because it is only under very specific circumstances and with restricted classes of nouns that the discourse constraints on VPE can be satisfied with a nominal antecedent.

Data availability/Supplementary files

Materials, raw data and statistical analyses can be found at: https://osf.io/qrhzk/?view_only=8fe240afc05045799b9049aa1e7cda9c The files are organized in three folders for Experiment 1, Experiment 2, and Experiment 3. Each folder includes anonymized data for the experiment as well as the R-script used for the analyses and the experimental materials. The folder for Experiment 2 includes the scripts for the raw ratings as well as the residuals.

Ethics and consent

The experiments reported in this paper were run in 2013 and 2014 at a time when the Université Paris Cité (then Université Paris Diderot) did not have an ethics committee for social sciences. We did, however, follow all the guidelines with respect to voluntary consent and data storage that are currently required by our ethics committee.

Acknowledgements

Preliminary versions of this paper were presented at the "Identity in Ellipsis" conference in Leiden on Sept. 20 2013 and at the "Linguistic Evidence Conference" in Tübingen on Feb. 14 2014. Some of the data were presented at the "Structure and Evidence in Linguistics" workshop in honor of Ivan Sag, in Stanford on April 29 2013. We would like to thank participants for comments. We would also like to thank Jonathan Ginzburg for discussions of QUD and concealed questions, Anne Jugnet for help with constructing Alt– contexts, Laura Kertz for discussions of mismatches, Corey Cusimano for setting up the experiments on Ibex and running them on Mechanical Turk, and for preliminary statistical analyses, Camille Dejarnett, Meg Grant and Gabriel Flambard for comments on the experimental materials and discussion. We thank Peter Culicover, Anne Jugnet, Laura Kertz, Jason Merchant, Geoff Pullum and Jeff Runner for detailed comments on a preliminary version of the manuscript. We would also like to thank the three anonymous reviewers, whose insightful comments helped us improve the paper both with respect to content and clarity of presentation. This paper is dedicated to the memory of Ivan Sag, a special friend, and one of the great pioneers in the study of ellipsis.

Competing interests

The authors have no competing interests to declare.

References

Aelbrecht, Lobke & Harwood, William. 2019. Predicate ellipsis. In Craenenbroek, Jeroen van & Temmerman, Tanja (eds.), *The Oxford handbook of ellipsis*, 504–525. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780198712398.013.23

AnderBois, Scott. 2014. The semantics of sluicing: Beyond truth conditions. *Language* 90. 887–926. DOI: https://doi.org/10.1353/lan.2014.0110

Ariel, Mira. 1990. *Accessing noun phrase antecedents*. London: Routledge.

Arregui, Ana & Clifton, Charles, Jr. & Frazier, Lyn & Moulton, Keir. 2006. Processing elided verb phrases with flawed antecedents: The recycling hypothesis. *Journal of Memory and Language* 55. 232–246. DOI: https://doi.org/10.1016/j.jml.2006.02.005

Barker, Chris. 2016. Why relational nominals make good concealed questions. *Lingua* 182. 12–29. DOI: https://doi.org/10.1016/j.lingua.2016.01.002

Baker, Carl L. 1968. *Indirect questions in English*. Urbana-Champaign, IL: University of Illinois dissertation.

Barr, Dale J. & Levy, Roger & Scheepers, Christoph & Tily, Harry J. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68. 255–278. DOI: https://doi.org/10.1016/j.jml.2012.11.001

Bélanger, Suzanne Michelle. 2014. *Regeneration in recall and verb phrase ellipsis*. Toronto: University of Toronto dissertation.

Bernardy, Jean-Philippe & Lappin, Shalom & Lau, Jey Han. 2018. The influence of context on sentence acceptability judgments. In *Proceedings of the 56th annual meeting of the association for computational linguistics*, 456–461. Melbourne: Association for Computational Linguistics. DOI: https://doi.org/10.18653/v1/P18-2073

Birner, Betty J. & Ward, Gregory. 1998. *Information status and noncanonical word order in English*. Amsterdam: Benjamins. DOI: https://doi.org/10.1075/slcs.40

Bürkner, Paul-Christian. 2017. brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software* 80(1). 1–28. DOI: https://doi.org/10.18637/jss.v080.i01

Clifton, Charles, Jr. & Frazier, Lyn. 2010. Imperfect ellipsis: Antecedents beyond syntax. *Syntax* 13. 279–297. DOI: https://doi.org/10.1111/j.1467-9612.2010.00142.x

Colonna, Saveria & Schimke, Sarah & Hemforth, Barbara. 2012. Information structure effects on anaphora resolution in German and French: A crosslinguistic study of pronoun resolution. *Linguistics* 50. 991–1013. DOI: https://doi.org/10.1515/ling-2012-0031

Cornish, Francis. 1999. *Anaphora, discourse, and understanding: Evidence from French and English*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780198236481.001.0001

Culicover, Peter W. & Jackendoff, Ray S. 2005. *Simpler syntax*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199271092.001.0001

Culicover, Peter W. & Jackendoff, Ray S. 2012. Same-except: A domain-general cognitive relation and how language expresses it. *Language* 88(2). 305–340. DOI: https://doi.org/10.1353/lan.2012.0031

Dalrymple, Mary. 2005. Against reconstruction in ellipsis. In Elugardo, Reinaldo & Stainton, Robert J. (eds.), *Ellipsis and nonsentential speech*, 31–55. Dordrecht: Springer. DOI: https://doi.org/10.1007/1-4020-2301-4_2

Dalrymple, Mary & Shieber, Stuart & Pereira, Fernando. 1991. Ellipsis and higher-order unification. *Linguistics and Philosophy* 14(4). 399–452. DOI: https://doi.org/10.1007/BF00630923

Davies, Mark. 2008–. The Corpus of Contemporary American English (COCA). https://www.english-corpora.org/coca/.

De la Fuente, Israel & Hemforth, Barbara. 2013. Effects of clefting and left-dislocation on subject and object pronoun resolution in spanish. In Cabrelli Amaro, Jennifer & Lord, Gillian & Prada Pérez, Ana de & Aaron, Jessi Elana (eds.), *Selected proceedings of the 16th Hispanic linguistics symposium*, 27–45. Somerville, MA: Cascadilla Proceedings Project.

Drummond, Alex. 2014. *Ibex farm*. http://spellout.net/latest_ibex_manual.pdf.

Dubey, Amit & Sturt, Patrick & Keller, Frank. 2005. Parallelism in coordination as an instance of syntactic priming: Evidence from corpus-based modeling. In *Hlt '05 proceedings of the conference on human langugage technology and empirical methods in natural language processing*, 827–834. Stroudsburg, PA: Association for Computational Linguistics. DOI: https://doi.org/10.3115/1220575.1220679

Elbourne, Paul. 2008. Ellipsis sites as definite descriptions. *Linguistic Inquiry* 39. 191–220. DOI: https://doi.org/10.1162/ling.2008.39.2.191

Elliott, Patrick D. & Nicolae, Andreea C. & Sudo, Yasutada. 2014. The sticky reading: VP ellipsis without parallel binding. In Snider, Todd & D'Antonio, Sarah & Weigand, Mia (eds.), *Proceedings of salt 24*. 640–655. New York: LSA. DOI: https://doi.org/10.3765/salt.v24i0.3639

Fiengo, Robert & May, Robert. 1994. *Indices and identity*. Cambridge, MA: MIT Press.

Flambard, Gabriel. 2018. *English VP anaphors: do it, do this, do that*. Paris: Université Paris Diderot dissertation.

Ford, Marilyn & Bresnan, Joan. 2013. They whispered me the answer. In King, Tracy Holloway & de Paiva, Valeria (eds.), *From quirky case to representing space*, 95–107. Stanford: CSLI Publications.

Frana, Ilaria. 2010. *Concealed questions. in search of answers*. Amherst, MA: University of Massachussetts dissertation.

Frana, Ilaria. 2017. *Concealed questions*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199670925.001.0001

Frana, Ilaria. 2020. Concealed questions. In Gutzmann, D. & Matthewson, L. & Meier, C. & Rullmann, H. & Zimmermann, T. E. (eds.), *Wiley Blackwell companion to semantics*, Hoboken, N.J.: Wiley. DOI: https://doi.org/10.1002/9781118788516.sem040

Frazier, Lyn. 2013. A recycling approach to processing ellipsis. In Cheng, Lisa Lai-Shen & Corver, Norbert (eds.), *Diagnosing syntax*, 485–501. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199602490.003.0024

Frazier, Lyn & Clifton, Charles, Jr. 2001. Parsing coordinates and ellipsis: Copy *α*. *Syntax* 4. 1–22. DOI: https://doi.org/10.1111/1467-9612.00034

Frazier, Lyn & Clifton, Charles, Jr. & Carlson, Katy. 2007. Focus and VP ellipsis. *Language and Speech* 50. 1–21. DOI: https://doi.org/10.1177/00238309070500010101

Frazier, Lyn & Duff, John. 2019. Repair or accommodation? split antecedent ellipsis and the limits of repair. *Glossa* to appear. DOI: https://doi.org/10.5334/gjgl.728

Frazier, Lyn & Munn, Alan & Clifton, Charles, Jr. 2000. Processing coordinate structures. *Journal of Psycholinguistic Research* 29. 343–370. DOI: https://doi.org/10.1023/A:1005156427600

Frazier, Lyn & Taft, Lori & Roeper, Tom & Clifton, Charles, Jr. & Ehrlich, Kate. 1984. Parallel structure: A source of facilitation in sentence comprehension. *Memory and Cognition* 12. 421–430. DOI: https://doi.org/10.3758/BF03198303

Fu, Jingqi & Roeper, Thomas & Borer, Hagit. 2001. The VP within process nominals: Evidence from adverbs and the VP anaphor *Do-So*. *Natural Language and Linguistic Theory* 19. 549–582. DOI: https://doi.org/10.1023/A:1010654105760

Geiger, Jeffrey & Xiang, Ming. 2021. At the syntax-discourse interface: Verb phrase ellipis interpretation in context. *Language* 97. e89–e110. DOI: https://doi.org/10.1353/lan.2021.0010

Ginzburg, Jonathan. 2012. *The interactive stance*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199697922.001.0001

Ginzburg, Jonathan & Miller, Philip. 2019. Ellipsis in Head-Driven Phrase Structure Grammar. In Craenenbroek, Jeroen van & Temmerman, Tanja (eds.), *The Oxford handbook of ellipsis*, 75–121. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780198712398.013.4

Goldberg, Adele E. 2006. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Grant, Margaret & Clifton, Charles & Frazier, Lyn. 2012. The role of non-actuality implicatures in processing elided constituents. *Journal of Memory and Language* 66. 326–343. DOI: https://doi.org/10.1016/j.jml.2011.09.003

Griffiths, James. 2019a. Beyond MaxElide: An investigation of A'-movement from elided phrases. *Linguistic Inquiry* 50(3). 571–607. DOI: https://doi.org/10.1162/ling_a_00317

Griffiths, James. 2019b. A q-based approach to clausal ellipsis: Deriving the preposition stranding and island sensitivity generalisations without movement. *Glossa* 4(1). 1–41. DOI: https://doi.org/10.5334/gjgl.653

Grimshaw, Jane. 1979. Complement selection and the lexicon. *Linguistic Inquiry* 10. 279–326.

Gundel, Jeanette & Hedberg, Nancy & Zacharski, Ron. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2). 274–307. DOI: https://doi.org/10.2307/416535

Hankamer, Jorge & Sag, Ivan A. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7. 391–426.

Hardt, Daniel. 1993. *Verb phrase ellipsis: Form, meaning, and processing*: University of Pennsylvania dissertation. Distributed as *IRCS Report*, 93–23.

Hardt, Daniel. 1999. Dynamic interpretation of verb phrase ellipsis. *Linguistics and Philosophy* 22. 185–219.

Harris, Jesse & Pylkkänen, Liina & McElree, Brian & Frisson, Steven. 2008. The cost of question concealment: Eye-tracking and MEG evidence. *Brain and Language* 107. 44–61. DOI: https://doi.org/10.1016/j.bandl.2007.09.001

Hemforth, Barbara & Konieczny, Lars & Scheepers, Christoph & Colonna, Saveria & Schimke, Sarah & Baumann, Peter & Pynte, Joël. 2010. Language specific preferences in anaphor resolution:

Exposure or gricean maxims? In Kruijff-Korbayová, Ivana & Steedman, Mark (eds.), *Proceedings of the 32nd annual conference of the cognitive science society*. Portland, USA.

Hofmeister, Philip & Casasanto, Laura Staum & Sag, Ivan A. 2014. Processing effects in linguistic judgment data: (super-)additivity and reading span scores. *Language and Cognition* 6. 111–145. DOI: https://doi.org/10.1017/langcog.2013.7

Hofmeister, Philip & Jaeger, T. Florian & Arnon, Inbal & Sag, Ivan A. & Snider, Neal. 2013. The source ambiguity problem: Distinguishing the effects of grammar and processing on acceptability judgments. *Language and Cognitive Processes* 28. 48–87. DOI: https://doi.org/10.1080/01690965.2011.572401

Huddleston, Rodney & Pullum, Geoffrey K., et al. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/9781316423530

Jacobson, Pauline. 2016. The short answer: Implications for direct compositionality (and vice versa). *Language* 92. 331–375. DOI: https://doi.org/10.1353/lan.2016.0038

Johnson, Kyle. 2001. What VP ellipsis can do, and what it can't, but not why. In Baltin, Mark & Collins, Chris (eds.), *The handbook of contemporary syntactic theory*, 439–479. Cambridge: Blackwell. DOI: https://doi.org/10.1002/9780470756416.ch14

Jugnet, Anne & Miller, Philip. 2024. Polar nouns and polar concealed questions. In Gardelle, Laure & Mignot, Elise & Neveux, Julie (eds.), *Nouns and the morphosyntax/semantics interface*, 233–268. London: Palgrave Macmillan. DOI: https://doi.org/10.1007/978-3-031-44561-3_9

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1. 3–44. DOI: https://doi.org/10.1007/BF00351935

Kehler, Andrew. 2000. Coherence and the resolution of ellipsis. *Linguistics and Philosophy* 23. 533–575. DOI: https://doi.org/10.1023/A:1005677819813

Kehler, Andrew. 2002. *Coherence, reference, and the theory of grammar*. Stanford: CSLI Publications.

Kehler, Andrew. 2015. On QUD-based licensing of strict and sloppy ambiguities. In *Proceedings of SALT 25*. 512–532. DOI: https://doi.org/10.3765/salt.v25i0.3071

Kehler, Andrew. 2019. Ellipsis and discourse. In Craenenbroek, Jeroen van & Temmerman, Tanja (eds.), *The Oxford handbook of ellipsis*, 314–341. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780198712398.013.13

Kertz, Laura. 2008. Focus structure and acceptability in verb phrase ellipsis. In Abner, Natasha & Bishop, Jason (eds.), *Proceedings of the 27th west coast conference on formal linguistics*. Cascadilla Proceedings Project.

Kertz, Laura. 2010. *Ellipsis reconsidered*. La Jolla, CA: University of California, San Diego dissertation.

Kertz, Laura. 2013. Verb phrase ellipsis: The view from information structure. *Language* 89. 390–428. DOI: https://doi.org/10.1353/lan.2013.0051

Keshet, Ezra. 2013. Sloppy identity unbound. In Snider, Todd (ed.), *Proceedings of salt 23*. 412–431. Santa Cruz, CA: LSA. DOI: https://doi.org/10.3765/salt.v23i0.2678

Kim, Christina S. & Kobele, Gregory M. & Runner, Jeffrey T. & Hale, John T. 2011. The acceptability cline in VP ellipsis. *Syntax* 14. 318–354. DOI: https://doi.org/10.1111/j.1467-9612.2011.00160.x

Kim, Christina S. & Runner, Jeffrey T. 2018. The division of labor in explanations of Verb Phrase ellipsis. *Linguistics and Philosophy* 41. 41–85. DOI: https://doi.org/10.1007/s10988-017-9220-0

Legendre, Géraldine & Grimshaw, Jane & Vikner, Sten. (eds.) 2001. *Optimality-theoretic syntax*. Cambridge, MA: MIT Press. DOI: https://doi.org/10.7551/mitpress/5161.001.0001

Levin, Nancy. 1986. *Main verb ellipsis in spoken English* (Outstanding Dissertations in Linguistics). New York: Garland.

Malt, Barbara C. 1985. The role of discourse structure in understanding anaphora. *Journal of Memory and Language* 24. 271–289. DOI: https://doi.org/10.1016/0749-596X(85)90028-2

Martin, Andrea E. & McElree, Brian. 2008. A content-addressable pointer mechanism underlies the comprehension of verb-phrase ellipsis. *Journal of Memory and Language* 58. 879–906. DOI: https://doi.org/10.1016/j.jml.2007.06.010

Mauner, Gail & Tanenhaus, Michael K. & Carlson, Greg N. 1995. A note on parallelism effects in processing deep and surface verb-phrase anaphora. *Language and Cognitive Processes* 10. 1–12. DOI: https://doi.org/10.1080/01690969508407085

Merchant, Jason. 2001. *The syntax of silence. Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oso/9780199243730.001.0001

Merchant, Jason. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27. 661–738. DOI: https://doi.org/10.1007/s10988-005-7378-3

Merchant, Jason. 2008. An asymmetry in voice mismatches in VP ellipsis and pseudogapping. *Linguistic Inquiry* 39. 169–179. DOI: https://doi.org/10.1162/ling.2008.39.1.169

Merchant, Jason. 2013. Voice and ellipsis. *Linguistic Inquiry* 44. 77–108. DOI: https://doi.org/10.1162/LING_a_00120

Miller, Philip. 2011. The choice between verbal anaphors in discourse. In Hendrickx, I. & Devi, S. Lalitha & Branco, A. & Mitkov, R. (eds.), *Anaphora processing and applications: 8th discourse anaphora and anaphor resolution colloquium, daarc 2011*, vol. 7099 (Lecture Notes in Artificial Intelligence). 82–95. Berlin: Springer.

Miller, Philip. 2013. Usage preferences: The case of the English verbal anaphor *do so*. In Müller, Stefan (ed.), *Proceedings of the 20th international conference on head-driven phrase structure grammar* (HPSG Online Proceedings). 121–139. Stanford: CSLI Publications. DOI: https://doi.org/10.21248/hpsg.2013.7

Miller, Philip. 2014. A corpus study of pseudogapping and its theoretical consequences. In Piñón, Christopher (ed.), *Empirical issues in syntax and semantics*, vol. 10. 73–90. www.cssp.cnrs.fr/eiss10/eiss10.pdf.

Miller, Philip & Hemforth, Barbara & Amsili, Pascal & Flambard, Gabriel. 2020. Missing antecedents found. In *Proceedings of the lsa*, vol. 5. 822–834. DOI: https://doi.org/10.3765/plsa.v5i1.4795

Miller, Philip & Pullum, Geoffrey K. 2014. Exophoric VP ellipsis. In Hofmeister, Philip & Norcliffe, Elisabeth (eds.), *The core and the periphery: Data-driven perspectives on syntax inspired by Ivan A. Sag*, 5–32. Stanford, CA: CSLI Publications.

Murphy, Gregory L. 1985. Processes of understanding anaphora. *Journal of Memory and Language* 24. 290–303. DOI: https://doi.org/10.1016/0749-596X(85)90029-4

Murphy, Gregory L. 1990. Interpretation of verb phrase anaphora: Influences of task and syntactic context. *The Quarterly Journal of Experimental Psychology Section A* 42. 675–692. DOI: https://doi.org/10.1080/14640749008401244

Nathan, Lance E. 2006. *On the interpretation of concealed questions*. Cambridge,MA: Massachussetts Institute of Technology dissertation.

Pater, Joe. 2009. Weighted constraints in generative linguistics. *Cognitive Science* 33. 999–1035. DOI: https://doi.org/10.1111/j.1551-6709.2009.01047.x

Phillips, Colin & Parker, Dan. 2014. The psycholinguistics of ellipsis. *Lingua* 151. 78–95. DOI: https://doi.org/10.1016/j.lingua.2013.10.003

Poppels, Till. 2020. *Towards a referential theory of ellipsis*. San Diego, CA: UC San Diego dissertation.

Poppels, Till. 2022. Explaining ellipsis without identity. *The Linguistic Review* 39(3). DOI: https://doi.org/10.1515/tlr-2022-2091

Poppels, Till & Kehler, Andrew. 2018. Overcoming the identity crisis: Novel evidence for a referential theory of Verb Phrase Ellipsis. In *Proceedings of the 53rd annual meeting of the chicago linguistic society*, 403–417. Chicago, IL.

Poppels, Till & Kehler, Andrew. 2019. Reconsidering asymmetries in voice-mismatched Verb Phrase ellipsis. *Glossa* 4(1). DOI: https://doi.org/10.5334/gjgl.738

Pullum, Geoffrey K. 1991. English nominal gerund phrases as noun phrases with verb phrase heads. *Linguistics* 29. 763–799. DOI: https://doi.org/10.1515/ling.1991.29.5.763

R Development Core Team. 2009. R: a language and environment for statistical computing. http://www.r-project.org.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. In Yoon, J.-H. & Kathol, A. (eds.), *Ohio state university working papers in linguistics*, vol. 49. 91–136. Columbus, OH.

Roberts, Leah & Duffield, Nigel & Matsuo, Ayumi. 2013. Processing VP-ellipsis and VP-anaphora with structurally parallel and nonparallel antecedents: An eye-tracking study. *Language and Cognitive Processes* 28. 29–47. DOI: https://doi.org/10.1080/01690965.2012.676190

Rooth, Mats. 1992. Ellipsis redundancy and reduction redundancy. In Berman, Steve & Hestvik, Arild (eds.), *Proceedings of the Stuttgart ellipsis workshop*.

Sag, Ivan A. 1976. *Deletion and logical form*. Cambridge, MA: MIT dissertation.

Sag, Ivan A. 2012. Sign-based construction grammar: An informal synopsis. In Boas, Hans C. & Sag, Ivan A. (eds.), *Sign-based construction grammar*, 69–202. Stanford, CA: CSLI Publications.

Sag, Ivan A. & Hankamer, Jorge. 1984. Towards a theory of anaphoric processing. *Linguistics and Philosophy* 7. 325–345. DOI: https://doi.org/10.1007/BF00627709

Schachter, Paul. 1977. Does she or doesn't she? *Linguistic Inquiry* 8. 763–767.

Shuval, Noa & Hemforth, Barbara. 2008. Accessibility of negated constituents in reading and listening. *Intercultural Pragmatics* 5. 445–469. DOI: https://doi.org/10.1515/IPRG.2008.022

Souesme, Jean-Claude. 1985. *Do Something et ses diverses réalisations en anglais contemporain*: Université Paris 7 dissertation.

Stockwell, Richard. 2020. *Contrast and Verb Phrase ellipsis: Triviality, symmetry, and competition*. Los Angeles: UCLA dissertation.

Stockwell, Richard. 2022. Contrast and verb phrase ellipsis: The case of tautologous conditionals. *Natural Language Semantics.* DOI: https://doi.org/10.1007/s11050-022-09189-3

Stockwell, Richard. 2023. Ellipsis, contradiction and voice mismatch. Ulster University. DOI: https://doi.org/10.16995/glossa.8382

Tanenhaus, Michael K. & Carlson, Greg N. 1990. Comprehension and deep and surface verbphrase anaphors. *Language and Cognitive Processes* 5. 257–280. DOI: https://doi.org/10.1080/01690969008407064

van Craenenbroeck, Jeroen. 2017. VP-ellipsis. In Everaert, Martin & Riemsdijk, Henk van (eds.), *The Blackwell companion to syntax*, 1–35. Hoboken: Wiley-Blackwell 2nd edn. DOI: https://doi.org/10.1002/9781118358733.wbsyncom049

Webber, Bonnie L. 1979. *A formal approach to discourse anaphora* (Outstanding Dissertations in Linguistics). New York: Garland.

Weir, Andrew. 2014. *Fragments and clausal ellipsis*. Amherst, MA: University of Massachusetts, Amherst dissertation.

Winkler, Susanne. 2000. Silent copy and polarity forcus in VP ellipsis. In Schwabe, Kerstin & Zhang, Ning (eds.), *Ellipsis in conjunction*, 221–246. Tübingen: Niemeyer. DOI: https://doi.org/10.1515/9783110952155-012

Winkler, Susanne. 2016. Ellipsis and information structure. In Féry, Caroline & Ishihara, Shinichiro (eds.), *The Oxford handbook of information structure*, 359–382. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780199642670.013.31

Xiang, Ming & Grove, Julian & Merchant, Jason. 2019. Structural priming in production through 'silence': an investigation of Verb Phrase ellipsis and null complement anaphora. *Glossa* 4(1). DOI: https://doi.org/10.5334/gjgl.726