

Interpretations of VP anaphora through reference to salient events

Kanan Luce, Jeffrey Geiger, Christopher Kennedy & Ming Xiang*

Abstract. Previous work has largely agreed that the verbal anaphors *do it/this/that* can freely be resolved with respect to information in either the linguistic or nonlinguistic context, whereas there has been more debate regarding the resolution strategy for verb phrase ellipsis. We present an experiment that confirms the intuition that interpretations of *do that* are highly sensitive to information in the nonlinguistic context. Comparison to a prior experiment on verb phrase ellipsis using the same paradigm reveals that ellipsis is much less sensitive to nonlinguistic manipulations, suggesting that it is not resolved solely with respect to salient discourse events, and that there is a separate linguistic identity requirement for ellipsis.

Keywords. verb phrase anaphora; do that; resolution; nonlinguistic context; antecedent; numerals; verb phrase ellipsis

1. Background. It has long been recognized that anaphoric processes differ with respect to the extent to which they can receive their interpretation from information available in the nonlinguistic context. Hankamer and Sag (1976) proposed a division between “deep anaphors”, like *do it*, which could receive their interpretation either from a linguistic antecedent or salient information in the nonlinguistic context, and “surface anaphors”, like verb phrase ellipsis (VPE), which they claimed could only get their interpretations from prior linguistic content. This distinction was demonstrated using sentences like those shown in (1) and (2).

- (1) Hankamer: I’m going to stuff this ball through this hoop.
 Sag: It’s not clear that you can do it.
 Sag: It’s not clear that you can ∅.
- (2) [Hankamer attempts to stuff a 9-inch ball through a 6-inch hoop.]
 Sag: It’s not clear that you can do it.
 Sag: #It’s not clear that you can ∅.

This formulation is functionally adequate for a number of research goals, such as the investigation of the licensing of formal mismatches between VPE sites and their linguistic antecedents (e.g., Dalrymple, et al. 1991, Hardt 1993, Fiengo and May 1994, Merchant 2001), and has largely been accepted wholesale in the literature. However, periodic observations of the apparent possibility of exophoric (i.e., antecedentless) VPE have called into question the cut-and-dry nature of the deep/surface divide (Schachter 1977, Merchant 2004, Miller and Pullum 2013).

These papers highlight naturally occurring instances of exophoric VPE such as that shown in (3). Although such exophoric ellipses seem to be quite rare, those that are attested are suggested to be relatively natural sounding and easily interpretable.

* Authors: Kanan Luce, University of California, Berkeley (kanan.luce@berkeley.edu); Jeffrey Geiger, University of Chicago (jgeiger@uchicago.edu); Christopher Kennedy, University of Chicago (ck@uchicago.edu) & Ming Xiang, University of Chicago (mxiang@uchicago.edu).

- (3) The aisles at the Lakewood Wal-Mart are surprisingly packed at 11 p.m. “Can we \emptyset ? Can we \emptyset ?” Vanessa tugs at her mother, pointing to a rack of “Lady and the Tramp” DVDs. Diaz shrugs. OK.

There has been some disagreement regarding how to account for exophoric VPE and what its relationship is to canonical cases of ellipsis in the context of an overt linguistic antecedent. Hankamer (1978) dismissed individual examples of exophoric VPE on various grounds, suggesting that some were simply idiomatically acceptable, while others were not counterexamples to Hankamer and Sag’s proposal because of their illocutionary force. Merchant (2004), who acknowledges the acceptability of some exophoric VPE constructions, similarly claims that they do not argue against a strict linguistic identity account for VPE interpretation. Merchant’s claim is that canonical VPE, in the presence of a linguistic antecedent, is resolved under strictly linguistic identity, and that exophoric ellipsis is not “true” VPE. Instead, Merchant proposes that hearers of exophoric VPE accommodate that the elided phrase is *do it*, an anaphor that can readily receive its interpretation from the nonlinguistic context in a way that canonical VPE cannot.

Miller and Pullum (2013), by contrast, claim that VPE is licensed with respect to a broader discourse context consisting of both linguistic and nonlinguistic information. Under their proposal, VPE is possible (and, by extension, interpretable) when this discourse context makes a particular set of alternative propositions highly salient, and the purpose of the elliptical utterance is to pick out one of the alternative propositions as being the case. Exophoric VPE is proposed to be possible under certain restricted discourse conditions, when a salient alternative set can be accommodated based on the nonlinguistic context and this accommodation is triggered by elliptical utterances of a certain form (e.g., a request for permission). Crucially, then, VPE interpretations are not generated under identity with a linguistic antecedent, but rather with respect to the information that is most salient in the discourse, broadly construed. Miller and Pullum claim that the illusion of a linguistic identity requirement for VPE comes from the fact that linguistic information is simply much more salient than nonlinguistic information, so the nonlinguistic context cannot license VPE very often, nor can the contribution of the nonlinguistic context to a VPE construction’s interpretation overwhelm the contribution of any linguistic antecedent.

These two models of VPE interpretation differ in the predictions they make for cases where the nonlinguistic context and the linguistic antecedent strongly promote the inclusion of different information in the interpretation of a VPE construction. According to Merchant’s proposal, information from the nonlinguistic context should be ignored when interpreting VPE accompanied by a linguistic antecedent, as canonical VPE interpretation is purely linguistic, and nonlinguistic information is only considered in cases of “defective” exophoric VPE. Miller and Pullum’s program, by contrast, suggests that information from the nonlinguistic context should be considered even when there is a useable linguistic antecedent, since the alternative set that is available in the broad discourse context is a function of the information available in both the nonlinguistic and linguistic modalities. Highly salient information in the nonlinguistic context should be able to affect the strength of various VPE interpretations, even if this effect is subordinate to that of the information from the linguistic antecedent.

Geiger and Xiang (2017) tested these two models using an experimental paradigm that investigated the available interpretations for VPE utterances while varying the level of support for certain interpretation-site content in both the nonlinguistic context and linguistic antecedent. Their results were more consistent with Miller and Pullum’s proposal insofar as the experiment

revealed that manipulation of the nonlinguistic context can influence the interpretation of a VPE site, even in the presence of a linguistic antecedent, although antecedent-identical VPE interpretations are always preferred. The results were incompatible with a model of VPE interpretation that gives no role to the nonlinguistic context in determining VPE interpretations in the presence of a linguistic antecedent.

However, it is not clear that Geiger and Xiang's results actually confirm Miller and Pullum's proposal. They clearly indicate a role for nonlinguistic information in VPE interpretation in the presence of a linguistic antecedent, but without a baseline of comparison that indicates the general discourse salience of the alternative interpretations, it is impossible to say whether the interpretations in Geiger and Xiang's experiment were completely determined by their discourse salience. In order to answer this question, the data for VPE must be compared with the data for another anaphor whose interpretations are known to be more or less deterministically generated by their salience in a broadly construed discourse context.

The goal of this paper is to extend Geiger and Xiang's experimental paradigm to investigate the relative contributions of nonlinguistic and linguistic information to the interpretation of the verbal anaphor *do that*. The anaphors *do it/do this/do that* are generally acknowledged as freely interpretable with respect to information available in either the nonlinguistic context or the linguistic context (Hankamer and Sag 1976, Miller 2011, Miller and Pullum 2013). This ability is not surprising given the parallel between these anaphors and, for example, demonstratives, which likewise are relatively easily interpreted with respect to salient information in the nonlinguistic context (Gundel et al. 1993).

The benefit of extending the paradigm to *do that* is twofold. First, the experiment represents a methodical investigation of a construction that has previously been described primarily in intuitive terms. The experiment should confirm or deny previous intuitions about the interpretability of *do that* given salient information exclusively in the linguistic or exclusively in the nonlinguistic modality, as well as provide new information about the relative role these information sources play in determining the anaphor's interpretation when they do not agree with one another on what information should be included. Second, insofar as the experiment verifies that *do that* can readily be resolved with respect to a salient nonlinguistic or linguistic event, the data for *do that* will serve as a baseline of comparison to determine whether Miller and Pullum's model of VPE resolution, in which the available interpretations are exhaustively determined by discourse salience when the preconditions on exophoric VPE use are met, is correct.

2. Experiment. The goal of the experiment presented here was to explore the available interpretations of the verbal anaphor *do that* when the linguistic antecedent and nonlinguistic context make different information available for inclusion in the interpretation. The experiment was an exact replication of Experiment 1 in Geiger and Xiang (2017), except that the critical verb phrase ellipses from that study were replaced with the verbal anaphor *do that*. In this experimental paradigm, subjects view nonlinguistic contexts and linguistic antecedents that vary with respect to how strongly they evoke certain numeral information, then rate interpretations of critical sentences that include or exclude the numeral from the anaphor's interpretation. Subjects' relative preference for including the numeral in the anaphor's interpretation reveals how strongly the nonlinguistic context and linguistic antecedent each contribute to the interpretation.

2.1. DESIGN & METHODS. The experiment had a 3 (nonlinguistic context) x 3 (linguistic antecedent) x 2 (interpretation prompt) design, for a total of 18 conditions. For each item, subjects

viewed a scenario title, a nonlinguistic context that was presented as a comic strip, and a visually presented discourse between two characters in the comic strip. The discourse consisted of an optional antecedent utterance (as “no antecedent” was a level for this factor), plus a critical reply from the other character that contained the *do that* anaphor. Beneath the discourse, subjects saw a prompt that provided one possible interpretation of the anaphor-containing utterance, and were asked to rate it from 1 to 7 to indicate how likely they thought it was to be the intended interpretation of the utterance, where 1 was the least likely. After the subjects entered their response, the next scenario appeared on the screen. A screenshot of a sample scenario is given in Figure 1.



Figure 1: Sample item showing experimental layout.

The experiment was conducted on the Amazon Mechanical Turk platform. Subjects were paid for their participation, which typically took less than 15 minutes. They were first familiarized with the task with two practice items, then viewed 6 critical scenarios, each in different conditions, plus 10 fillers, in a random order.

2.2. MATERIALS. The levels for the nonlinguistic context manipulation were *unavailable*, in which there was no recoverable numeral information for the referent of interest; *available*, in which numeral information about the referent was made recoverable, but not highly salient, by having a character interact with the set of objects as a group; and *salient*, in which the same numeral information was made highly salient by having a character interact with the objects one by one while counting them.

The levels for the linguistic antecedent were *exophoric*, in which case there was no antecedent; *unmodified*, where the antecedent utterance contained no numeral; and *modified*, where the antecedent utterance contained the numeral that was optionally evoked in the corresponding nonlinguistic contexts.

For every item, the critical anaphor-containing utterance featured the *do that* anaphor under negation, and in particular, a comment about the impossibility or undesirability of taking the action suggested by the linguistic antecedent. The choice of anaphor contrasts with Geiger and Xiang (2017), where the critical utterance was a verb phrase elliptical comment about the action’s impossibility or undesirability (e.g., “We can’t.”), which is argued by Miller and Pullum to trig-

ger accommodation of a salient set of alternatives from the nonlinguistic context and thus license exophoric VPE.

One of the two possible interpretation prompts that followed the anaphor-containing reply was *polar*. This interpretation arises when the anaphor site is interpreted as not containing a numeral (e.g., “buy candy bars”). Since the anaphor site is in the scope of negation, the only interpretation when there is no numeral is that the entire proposed action is not possible or advisable. The other interpretation is *scalar*. This interpretation corresponds to the presence of a numeral in the anaphor site (e.g., “buy five candy bars”). Under negation, this is compatible with a reading under which the negation applies to the cardinality of the relevant set of objects, and not to the proposed action itself. In other words, a possible interpretation with a numeral in the anaphor site is that the proposed action is possible, but only with a set of objects smaller than that previously proposed.

Sample levels for one critical scenario are shown in Table 1, while the nonlinguistic context comic strips for the same scenario are shown in Figure 2.

Nonlinguistic Context	Antecedent	Reply	VPE Interpretation
<i>Unavailable</i> : Father and son stand in grocery store aisle near candy bars.	<i>Exophoric</i> : [no antecedent]	Father: We can't do that.	<i>Polar</i> : On a scale from 1 to 7, where 1 is the least likely and 7 is the most likely, how likely do you think it is that the father meant: We can't buy any candy bars.
<i>Available</i> : Son places five candy bars in cart at one time.	<i>Unmodified</i> : Son: I want to buy candy bars!	(<i>cf. Geiger & Xiang</i>) Father: We can't.	<i>Scalar</i> : ...We can't buy five candy bars, but maybe we could buy fewer.
<i>Salient</i> : Son conspicuously places five candy bars in cart one at a time.	<i>Modified</i> : Son: I want to buy five candy bars!		

Table 1: Experiment factors and sample levels.

2.3. PARTICIPANTS. The participants were Amazon Mechanical Turk users living in the United States and aged between 18 and 49 (mean = 29). Only participants who self-reported as native speakers of English were included in the analysis. A total of 89 subjects were included.

2.4. PREDICTIONS. Of interest in this study is subjects' relative preference for including or excluding a numeral from their interpretation of the *do that* anaphor given varying degrees of support for the numeral in both the nonlinguistic context and the linguistic antecedent. Given that *do that* has been intuitively described as freely deriving its interpretation from either the linguistic or nonlinguistic context, the results of the experiment should show that the interpretation is highly sensitive to manipulations in either the linguistic or nonlinguistic context when isolated from the effect of the other. For instance, in the exophoric conditions, where there is no linguistic antecedent, increasing numeral salience in the nonlinguistic context should be accompanied by increased ratings for the scalar (numeral-modified) interpretation, and decreased ratings for the polar interpretation. Likewise, with the unavailable nonlinguistic context, which makes a minimal contribution to the discourse in the sense that it contains no salient event, including a numeral in the linguistic antecedent (modified) should boost the scalar interpretation and suppress the polar interpretation relative to the unmodified antecedent, which does not contain a numeral.



Figure 2: Sample nonlinguistic contexts. Top: unavailable. Middle: available. Bottom: salient.

What is less clear is the predicted pattern of results when the nonlinguistic context and the linguistic antecedent disagree on what information should be included in the anaphor interpretation. For instance, in the conditions with a salient context and an unmodified antecedent, the nonlinguistic context makes numeral information about the target referent highly salient, but the absence of a numeral in the linguistic antecedent suggests that no numeral should be included in the anaphor’s interpretation. There are multiple conceivable strategies for interpreting *do that* in such a situation.

According to one strategy, the linguistic antecedent is “privileged” in the anaphor’s interpretation, either because linguistic identity is the anaphor’s “primary” locus of resolution, or because resolution is primarily driven by discourse salience, and linguistically mentioned material is inherently much more salient than nonlinguistic material (Miller and Pullum 2013). Under the conditions described above, this would cause the linguistic context to “win”, meaning subjects would favor the polar interpretation over the scalar interpretation.

Alternately, it may be the case that salient information from any source, linguistic or non-linguistic, can be included in an anaphor’s interpretation. In this case, it is not clear that the polar interpretation would “win” so readily given an unmodified linguistic antecedent and salient non-linguistic context. Instead, the ratings for the two interpretations might be closer to one another, or the ratings for the scalar interpretation might even exceed the ratings for the polar interpretation depending on *do that*’s relative sensitivity to linguistic vs. nonlinguistic information.

2.5. RESULTS. The mean ratings and standard errors for the polar and scalar interpretations in each combination of nonlinguistic context and linguistic antecedent are shown in Table 2. These results are also plotted in Figure 3.

Interpretation	Context	Antecedent		
		Exophoric	Unmodified	Modified
Polar	Unavailable	4.94 (.273)	5.12 (.378)	3.81 (.370)
	Available	3.42 (.356)	5.17 (.333)	3.48 (.356)
	Salient	3.42 (.356)	4.13 (.387)	2.79 (.314)
Scalar	Unavailable	3.79 (.408)	3.09 (.349)	5.38 (.319)
	Available	5.11 (.318)	3.73 (.371)	6.11 (.235)
	Salient	5.23 (.243)	5.20 (.300)	5.15 (.286)

Table 2: Mean rating and standard error by condition.

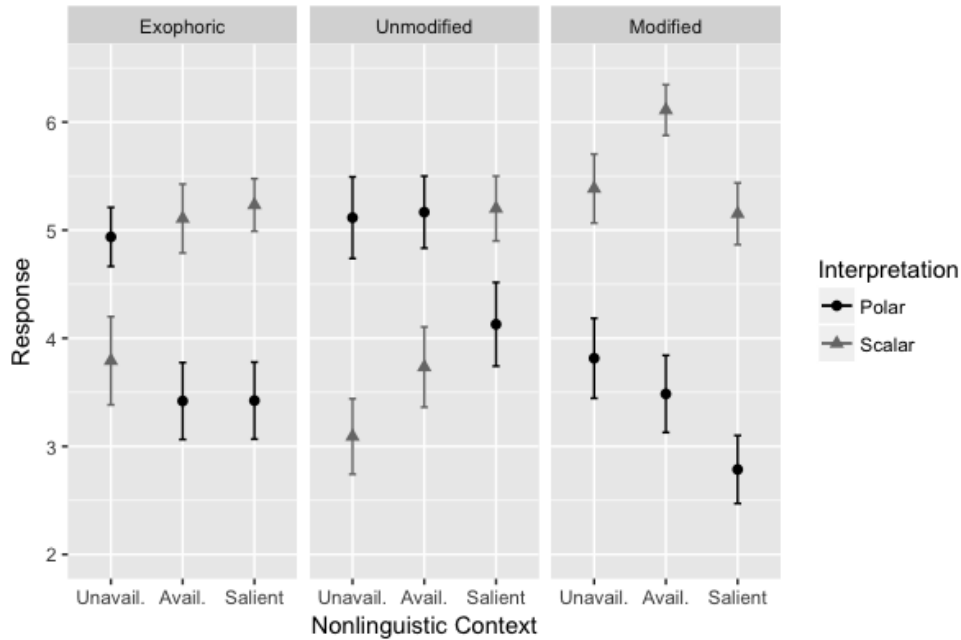


Figure 3: Mean rating by experimental condition. Error bars: standard error.

2.6. ANALYSIS. The response data were fit to a linear mixed-effects regression model with fixed effects of nonlinguistic context, linguistic antecedent, and interpretation prompt, two-way interactions for each pair of these factors, and a three-way interaction of all three factors, plus random effects for subject and scenario. There was a significant three-way interaction between nonlinguistic context, linguistic antecedent, and interpretation prompt ($p < .01$), so further analysis was carried out on subsets of the data separated by linguistic antecedent type to examine the effect of manipulations in the nonlinguistic context with a constant linguistic antecedent. Each sub-analysis consisted of a linear mixed-effects regression model with fixed effects of nonlinguistic context and interpretation prompt, their interaction, and random effects for subject and scenario.

For the exophoric conditions, there was a significant interaction between nonlinguistic context type and the interpretation rated ($p < .001$), suggesting that the effect of the nonlinguistic manipulation on ratings differed depending on the interpretation being rated. Paired comparisons suggested that the ratings for the scalar reading increased and the ratings for the polar reading decreased as numeral information was introduced in the nonlinguistic context, as the ratings for each interpretation in the unavailable context were significantly different for the ratings of the same reading in the available and salient contexts (all p 's $< .01$). Ratings of the same reading in

the available vs. salient context were not significantly different from one another ($p's > .7$). The ratings for the polar and scalar interpretations were significantly different from one another in all three nonlinguistic contexts (all $p's < .05$).

For the conditions with an unmodified linguistic antecedent, there was also a significant interaction between nonlinguistic context and interpretation prompt ($p < .001$). Paired comparisons suggested that the ratings for the polar interpretation were lower in the salient context than in the other two contexts (unavailable-salient $p < .05$; available-salient $p \approx .05$). Likewise, the ratings of the scalar interpretation in the salient context were significantly different from the ratings in the other two contexts ($p's < .001$). The ratings of either interpretation in the unavailable vs. available contexts were not significantly different from one another ($p's > .2$). The ratings for the two interpretations were significantly different from one another in all three nonlinguistic contexts (all $p's < .05$).

In the conditions with a modified linguistic antecedent, there was no significant interaction between nonlinguistic context type and the interpretation rated ($p > .2$). This suggests that the effect, if any, of manipulating the nonlinguistic context did not depend on the interpretation being rated. In these conditions, ratings were primarily determined solely by the reading being rated (fixed effect $p < .001$). No paired comparisons were performed in these conditions.

2.7. DISCUSSION. The results of the experiment confirmed previous intuitive judgments of *do that* as highly sensitive to both nonlinguistic and linguistic information. In the exophoric conditions, the results indicated that *do that* can readily be resolved with respect to a salient event in the nonlinguistic context when there is no linguistic antecedent. In the unavailable nonlinguistic context, the polar (numeral-unmodified) interpretation was significantly preferred over the scalar interpretation. This condition features minimal informational contributions from both the nonlinguistic context and the linguistic antecedent, so it is not surprising that subjects did not infer the presence of a numeral in the anaphor site. However, as the numeral information was made more available in the exophoric conditions via the available or salient contexts, subjects' preferences between the two interpretations flipped, with the scalar interpretation being significantly preferred over the polar interpretation. Thus, in the exophoric conditions, subjects appear to rely heavily on the information that is salient in the nonlinguistic context when constructing an interpretation for *do that*.

Likewise, the interpretation of *do that* is heavily dependent on the content of the linguistic antecedent. The patterns of results were distinct for each of the three antecedent types, indicating that the type of linguistic antecedent was critical in determining the ratings for each interpretation. Unsurprisingly, in the unavailable nonlinguistic context, which makes a relatively small contribution to the discourse, the polar interpretation was significantly preferred in the presence of the unmodified antecedent, while the scalar interpretation was preferred with a modified antecedent. In other words, the linguistic antecedent more or less deterministically predicted the preferred interpretation when the nonlinguistic context made a minimal contribution.

Perhaps more interesting are the conditions in which the nonlinguistic context and linguistic antecedent both make a highly salient contribution to the discourse. In the conditions with unmodified antecedents, the linguistic antecedent represents a highly salient event that omits numeral information. As mentioned above, in the absence of any salient nonlinguistic information supporting the inclusion of a numeral in the anaphor's interpretation, the polar interpretation is preferred. This pattern holds in the available nonlinguistic context as well, but when the salient

nonlinguistic context, which makes the numeral information highly salient, is present, subjects' preferences for the two interpretations flip, with the scalar interpretation significantly preferred over the polar interpretation. By contrast, the nonlinguistic context manipulation was not associated with a statistically reliable effect on ratings for the conditions with a modified linguistic antecedent, although there was a trend toward lower ratings for the polar interpretation with increasing numeral availability in the nonlinguistic context.

These results are interesting because they point to a nuanced relationship between nonlinguistic and linguistic information in the construction of interpretations for *do that*. It is clear that information from both sources is considered during the interpretation of the anaphor, as demonstrated by the effect of the nonlinguistic context manipulation on the ratings in the conditions with an unmodified linguistic antecedent. This argues against a model according to which resolution of *do that* under linguistic identity is highly privileged over its resolution with respect to a salient nonlinguistic event, as well as a model that suggests the interpretation is a function of general discourse salience, with linguistic material being inherently much more salient than nonlinguistic material. Indeed, the fact that the scalar interpretation of the *do that* sentence is preferred over the polar interpretation when there is a salient nonlinguistic context and unmodified linguistic antecedent shows that a highly salient nonlinguistic event can “outweigh” overt linguistic evidence (i.e., the omission of the numeral from the antecedent) that a particular piece of information is not important for the interpretation of an anaphor.

Rather, the results suggest that both the nonlinguistic and linguistic context in which *do that* are considered during its interpretation. In the presence of an unmodified linguistic antecedent, the polar interpretation is preferred, except when the numeral information is made highly available in the salient nonlinguistic context. In other words, subjects were willing to interpret *do that* in an antecedent-nonidentical way given strong nonlinguistic evidence that the numeral was relevant. That this switch occurred only in the salient context, whereas in the exophoric conditions the scalar interpretation was already rated better in the available context, suggests that the explicit lack of a numeral in the linguistic antecedent does make subjects more hesitant to accept the scalar interpretation, but that this effect does not completely overwhelm the contribution of very salient numeral information in the nonlinguistic context. In contrast, with a modified linguistic antecedent, the presence of a numeral in the linguistic antecedent appears to overwhelm all other factors, such that the scalar interpretation is uniformly preferred in all nonlinguistic contexts.

Thus, it seems safe to conclude that neither nonlinguistic nor linguistic information is “dominant” in the interpretation of *do that*. Salient numeral information in both the nonlinguistic and linguistic content can effectively promote the inclusion of a numeral in the anaphor's interpretation. It is worth noting that the effect of salient numeral information in the linguistic antecedent is somewhat more resilient than that of salient information in the nonlinguistic context, in the sense that the modified antecedent alone was sufficient to promote the scalar interpretation in all nonlinguistic contexts, whereas the presence of an unmodified antecedent somewhat counteracted the ability of salient nonlinguistic information to promote the scalar interpretation compared to the exophoric conditions. This may reflect subjects' intuitions about the commitments of the participants in the experimental discourse; by uttering the unmodified or modified linguistic antecedent, the character is committed to the numeral information being either irrelevant or relevant to the discourse, whereas the characters' mere presence in a particular nonlinguistic context requires no such commitment. Given that the utterer of the anaphoric sentence should be making a reply

that is relevant to the antecedent utterance, it is not surprising that subjects were somewhat more hesitant to select an interpretation that is not identical to the linguistic antecedent vs. one that is inconsistent with the numeral information in the nonlinguistic context. However, the experimental results clearly demonstrate that this hesitation can be overcome given a sufficiently salient event in the nonlinguistic context, and that information from both the nonlinguistic and linguistic context are considered during the interpretation of *do that*.

3. General discussion. In addition to investigating the relative effects of nonlinguistic and linguistic information on the interpretation of *do that*, this experiment represents an opportunity to compare the interpretation strategy for this anaphor with the strategy for verb phrase ellipsis, since the paradigm was identical to that used by Geiger and Xiang (2017) with VPE critical sentences. Geiger and Xiang’s results for the VPE experiment (their Experiment 1) are shown in Figure 4.

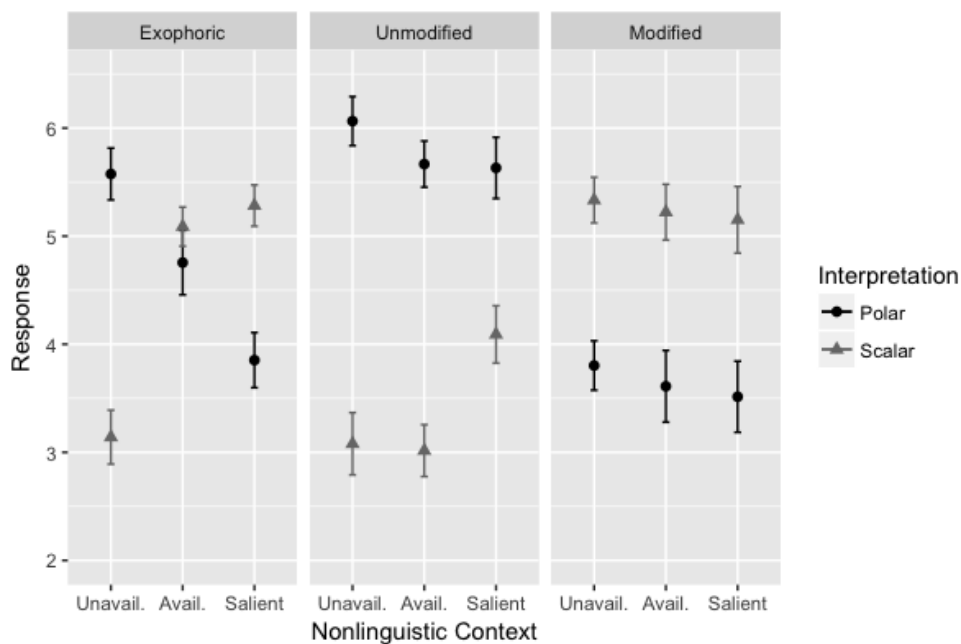


Figure 4: Geiger & Xiang’s (2017) results for VPE. Error bars: standard error.

The results for *do that* and VPE show some striking similarities, but there are also key distinctions that suggest that their interpretation strategies differ in critical ways. In the exophoric conditions, the scalar interpretation ratings increased and the polar interpretation ratings decreased as a function of nonlinguistic numeral availability in both experiments. However, this change was more pronounced in the *do that* experiment, where the scalar interpretation was significantly preferred over the polar interpretation even in the available context, whereas for VPE, the interpretations were rated approximately equally in the available context. This suggests that with no useable linguistic antecedent, the baseline sensitivity of *do that* to salient events in the nonlinguistic context is somewhat higher than that of VPE, since subjects were more willing to incorporate numeral information of medium nonlinguistic salience into the interpretation of *do that* than the interpretation of VPE.

For the conditions with an unmodified linguistic antecedent, the results for the two experiments were similar in kind, but not in magnitude. For *do that*, the (antecedent-identical) polar

interpretation was preferred over the scalar interpretation in the unavailable and available contexts, but when the salient context made the numeral information highly salient, the ratings for the scalar interpretation increased and the ratings for the polar interpretation decreased such that the scalar interpretation was actually preferred. The VPE experiment saw a similar boost for the scalar interpretation in only the salient context, but the antecedent-identical polar interpretation was uniformly preferred in all three nonlinguistic contexts. This difference suggests that unlike *do that*, the interpretation strategy for VPE does actually privilege identity with a linguistic antecedent in a significant way. This comparison argues against the proposal that VPE is interpreted with respect to a broad (nonlinguistic and linguistic) discourse context, where the appearance of a linguistic identity requirement is simply a result of a salience advantage for linguistic over nonlinguistic information. Instead, the role of nonlinguistic information in determining VPE interpretations in the presence of a linguistic antecedent is clearly more limited than for *do that* in a way that likely cannot be reduced to a simple salience gap.

The results of the two experiments were most similar in the conditions with a modified antecedent, where the presence of the numeral in the linguistic antecedent dominated the interpretation, and the scalar reading was preferred across the board. For both constructions, it seems that the inclusion of the numeral in the linguistic antecedent was sufficient to generate certainty that the reply anaphor would also include a numeral in its interpretation, perhaps because the linguistically evoked numeral (or the speaker's commitment to its relevance) was extremely salient, or, particularly in the case of VPE, because a linguistic identity requirement more or less mandates its inclusion in the anaphor's interpretation.

4. Conclusion. The experiment presented here represents a methodical investigation of the relevant contributions of nonlinguistic and linguistic information in the interpretation of the verbal anaphor *do that*. The results of the experiment confirmed prior intuitions that *do that* can freely be interpreted with respect to a salient event in either the nonlinguistic context (e.g., a salient physical event) or the linguistic context (e.g., an utterance). Moreover, the experiment revealed that in cases where salient information is available in both the nonlinguistic and linguistic context, both information sources are considered and can influence the interpretation of the anaphor. Linguistic information is somewhat privileged in these cases, perhaps because it is more salient or subjects are more willing to ascribe awareness of it to interlocutors because of the underlying communicative intent, but sufficiently salient nonlinguistic information can still override any preference for identity between the anaphor interpretation and the linguistic antecedent.

Further, since the experiment used the same paradigm as that used by Geiger and Xiang (2017) to investigate the interpretation of VPE, a preliminary comparison of the interpretation mechanisms for the two phenomena was made possible. This comparison suggested that *do that* was somewhat more sensitive to nonlinguistic manipulations than was VPE. More interestingly, the comparison argued against theories of VPE that suggest it is interpreted with respect to a broad discourse context, and that the available interpretations are a function of their relative discourse salience. Whereas highly salient nonlinguistic information could override a slight preference for linguistic identity for *do that*, the same was not true for VPE. While salient nonlinguistic information could boost consideration of an antecedent-nonidentical interpretation for VPE, the nonidentical interpretation was never preferred over, or even equal to, the identical interpretation. This suggests that the VPE interpretation mechanism does include a special role for linguistic identity that other anaphors, like *do that*, may not, although strictly linguistic accounts of VPE

interpretation are also insufficient insofar as they fail to account for the ability of nonlinguistic information to boost consideration of nonidentical interpretations.

The results of the two experiments show that the model of VPE interpretation must be more nuanced than the majority of prior, monolithic theories have allowed. Indeed, such models have recently been suggested. Merchant (2010, 2016) proposes the use of stochastic ranked decision trees to model the resolution of VPE constructions. In particular, the question “Is there a linguistic antecedent?” is proposed to rank highly in the tree, with different resolution strategies depending on the answer. This approach allows for resolution under strict linguistic identity if the answer is “yes”, but resolution using a different strategy if the answer is “no”. Interpreted in this light, Geiger and Xiang’s results showing a boost for antecedent-nonidentical interpretations in the presence of a useable antecedent could be modeled in terms of subjects’ uncertainty as to whether the antecedent was sufficiently recoverable to the discourse participants in context. In other words, the available ellipsis interpretations would not be a direct function of discourse salience, but rather a function, mediated by discourse salience, of subjects’ expectations about what resolution strategy the speaker of a VPE utterance expects the hearer to employ in its resolution.

Further work in this realm could focus on extending the investigation to include other anaphors, which would shed light on how various constructions are similar or different with respect to the relative contributions of nonlinguistic and linguistic information. Switching from an offline to an online experimental paradigm could reveal whether nonlinguistic and linguistic information are considered simultaneously during anaphor interpretation, or whether an initial interpretation is constructed on the basis of only one information source, and this is later revised in light of information that is made available by the other. Future research is particularly necessary on the interpretation mechanism for VPE, since Geiger and Xiang’s experiment and its comparison to the *do that* experiment presented here, taken together, argue against both a purely discourse-based, salience-driven account of VPE interpretation as well as an account based only on linguistic identity.

References

- Dalrymple, Mary, Stuart M. Sheiber & Fernando C. N. Pereira. 1994. Ellipsis and higher-order unification. *Linguistics and Philosophy* 14. 399-452. <https://doi.org/10.1007/BF00630923>.
- Fiengo, Robert & Robert May. 1994. *Indices and Identity*. Cambridge, MA: MIT Press.
- Geiger, Jeffrey & Ming Xiang. 2017. “Context can!”: Contextual accommodation in exophoric and anaphoric verb phrase ellipsis. *Proceedings of the North East Linguistics Society* 47(2). 35-44.
- Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2). 274-307. <https://doi.org/10.2307/416535>.
- Hankamer, Jorge. 1978. On the nontransformational derivation of some null VP anaphors. *Linguistic Inquiry* 9(1). 66-74.
- Hankamer, Jorge & Ivan Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7(3). 391-428.
- Hardt, Daniel. 1993. *Verb phrase ellipsis: Form, meaning, and processing*. Philadelphia, PA: University of Pennsylvania dissertation.

- Merchant, Jason. 2001. *The Syntax of Silence*. Oxford: Oxford University Press.
- Merchant, Jason. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27. 661-738. <https://doi.org/10.1007/s10988-005-7378-3>.
- Merchant, Jason. 2010. Three kinds of ellipsis. In Recanati, Francois, Isidora Stanoic & Nef-tali Villanueva (eds.) *Context-Dependence, Perspective, and Relativity*. Berlin: Walter de Gruyter. 141-192.
- Merchant, Jason. 2016. One size doesn't fit all: What fragments can tell us about anaphoricity, and why reductionist monolithic theories fail. Presentation at the Workshop on Fragments, Saarland University, Saarbrücken, Germany.
- Miller, Philip. 2011. The choice between verbal anaphors in discourse. *Anaphora Processing and Applications: 8th Discourse Anaphora and Anaphor Resolution Colloquium*. 82-95.
- Miller, Philip & Geoffrey K. Pullum. 2013. Exophoric VP ellipsis. In Hofmeister, Philip & Elisabeth Norcliffe (eds.) *The Core and the Periphery: Data-driven Perspectives on Syntax Inspired by Ivan A. Sag*. Chicago, IL: University of Chicago Press.
- Schachter, Paul. 1977. Does she or doesn't she? *Linguistic Inquiry* 8(4). 763-767.