

# Perspective-shifting with appositives and expressives

Jesse A. Harris · Christopher Potts

© Springer Science+Business Media B.V. 2010

**Abstract** Much earlier work claims that appositives and expressives are invariably speaker-oriented. These claims have recently been challenged, most extensively by Amaral et al. (*Linguist and Philos* 30(6): 707–749, 2007). We are convinced by this new evidence. The questions we address are (i) how widespread are non-speaker-oriented readings of appositives and expressives, and (ii) what are the underlying linguistic factors that make such readings available? We present two experiments and novel corpus work that bear directly on this issue. We find that non-speaker-oriented readings, while rare in actual language use, are systematic. We also find that non-speaker-oriented readings occur even outside of attitude predications, which leads us to favor an account based in pragmatically-mediated perspective shifting over one that relies on semantic binding by attitude predicates.

**Keywords** Appositives · Expressives · Perspective · Corpus pragmatics · Experimental pragmatics · Regression analysis

## 1 Introduction

Amaral et al. (2007) (henceforth ARS) is a critical review of Potts (2005). The authors articulate challenging new questions for researchers working on conversational dynamics, pragmatic inference, scope, and quantification. The present paper addresses phenomena at the confluence of these areas: the range of readings available for appositives and expressives, and the underlying semantic and pragmatic factors that produce those readings.

---

J. A. Harris  
University of Massachusetts, Amherst, MA 01003, USA  
e-mail: harris@linguist.umass.edu

C. Potts (✉)  
Stanford University, Stanford, CA 94305-21050, USA  
e-mail: cgpotts@stanford.edu

We say that a clause  $C$  with denotation  $p$  is *speaker-oriented* in utterance  $U$  if, and only if, in uttering  $U$ , the speaker expresses, with  $C$ , a commitment to  $p$ . For example, if I utter the clause *Ohio is the birthplace of aviation* with the intention of sincerely asserting it, then I express, in virtue of this utterance, my public commitment to the proposition that Ohio is the birthplace of aviation. Thus, *Ohio is the birthplace of aviation* is speaker-oriented in this case. We emphasize that speaker-orientation is a relation between an individual and a linguistic structure, but that the relationship depends not only on semantic denotations but also on utterances and their complex pragmatics.

Lakoff (1966), McCawley (1982, 1987, 1989), Asher (2000), Huddleston and Pullum (2002), Emonds (1976), Culicover (1992), Bach (1999), and Potts (2005) claim or assume that appositive relatives are always speaker-oriented, regardless of syntactic position, and Quang (1971), Cruse (1986), Aoun et al. (2001), Corazza (2005), and Potts (2005) take a similar position on epithets like *the jerk*.<sup>1</sup> These generalizations were first challenged by Wang et al. (2005) and Karttunen and Zaenen (2005). ARS present a variety of examples in which such clauses can be non-speaker-oriented when uttered (see also Potts 2005, 162, 2007). We are persuaded by this new evidence. We take it as a starting point for answering the following more specific questions about the semantics and pragmatics of these constructions:

- (1) a. How widespread are non-speaker-oriented readings of appositives and expressives?
- b. What are the underlying linguistic factors that make such readings available?

We think that the investigative strategy of reporting basic intuitions about individual cases has run its course in this area. More and different evidence is needed. To this end, we present two human-subjects experiments and some novel corpus work. Taken together, the results from this research support clear answers to the questions in (1). We find that non-speaker-oriented readings, while rare in actual language use, are systematic: experiment 1 involves contexts in which non-speaker-oriented readings of appositives are favored, and experiment 2 pinpoints one of the pragmatic factors that facilitate non-speaker-orientation for expressives. We also find that non-speaker-oriented readings occur even outside of attitude predications, which leads us to favor an account based in pragmatically-mediated perspective shifting (Potts 2007) over one that relies on semantic binding by attitude predicates (Schlenker 2003, 2007; Sauerland 2007). The results challenge Potts's (2005) naïve view of speaker-orientation, but they are consistent with the multidimensional theory of composition he develops.

The paper is structured as follows. Section 2 describes our empirical focus in greater detail and reviews previous evidence for non-speaker-oriented readings. Sections 3 and 4 present our human-subjects experiments, and Sect. 5 reports on our

<sup>1</sup> An exception must be made for direct quotation; Potts (2005), Bach (2006), Anand (2007).

corpus work. Finally, Sect. 6 draws some general conclusions from this evidence about the centrality of pragmatics in understanding these phenomena.

## 2 Appositives, expressives, and (non-)speaker orientation

Both Potts (2005) and ARS discuss many different kinds of appositive and expressive. In this paper, we restrict attention to just nominal appositives, as in (2a), appositive relatives, as in (2b), and nominal epithets like (2c).<sup>2</sup>

- (2) a. Lucille Gorman, an 84-year-old Chicago housewife, has become amazingly immune to stock-market jolts.  
[Treebank corpus; Marcus et al. 1999]
- b. uh, she starts a new job tomorrow, which should take her out of the house about four days a week.  
[Switchboard corpus; Godfrey and Holliman 1993]
- c. In traffic so heavy that there is no way for the jerk to pass, I might pull over, as if to look for a street number or name, (still ignoring the jerk) just to get the jerk off my tail.  
[20\_newsgroups corpus]

These constructions are ideal for present purposes because they have an extremely wide syntactic distribution. One can embed them in many different kinds of environments, and they are used in a wide variety of different discourse contexts to convey a wide variety of different messages.

It is uncontroversial that these expressions can scope out of the usual array of presupposition holes (interrogatives, negations, conditional antecedents, and modals; Karttunen 1973), in the sense that they can remain semantically unmodified by those operators even when in their syntactic scope. We illustrate briefly in (3) using conditionals; for additional examples and discussion, see Boër and Lycan (1976), Kaplan (1999), Aoun et al. (2001), Karttunen and Zaenen (2005), Potts (2005, to appear).

- (3) a. I think it would concern me even more if I had children, which I don't, [...] [Switchboard]
- b. If that bastard Kaplan was promoted, then the Regents acted foolishly. [Kaplan 1999]

In (3a), the appositive *which I don't* is syntactically inside the if-clause, but interpreting it there semantically would make this clause semantically contradictory, which is clearly not what the speaker intends. Rather, the appositive stands, unconditionalized, as a speaker commitment. Similarly, (3b) highlights, in Kaplan's (1999) words, the "failure of the conditional to conditionalize away expressive content".

<sup>2</sup> One must take care to distinguish these appositives from sliftings like *Ed, it seems, is a werewolf* (Ross 1973; Asher 2000; Wagner 2004) and predicative clauses like *As president, I would launch a mission to Mars*. Indeed, almost everything we say about nominal appositives and appositive relatives is false of these other constructions.

It is also uncontroversial that appositives and expressives can scope out of the complements of attitude predicates and verbs of saying, thereby becoming speaker-oriented even as the clausal material around them is attributed to the embedding subject (Quang 1971; Cruse 1986; Bach 1999; Aoun et al. 2001; Potts 2005; Taylor 2007). The following passage from the text of Aloni (2000) is useful for highlighting this exceptional scopal behavior for appositives:

- (4) In front of Ralph stand two women. Ralph believes that the woman on the left, who is smiling, is Bea, and the woman on the right, who is frowning, is Ann. As a matter of fact, exactly the opposite is the case.
- S1: 

Bea ☺	Ann ☺
-------	-------
- S2: 

Bea ☺	Ann ☹
-------	-------
- S3: 

Ann ☹	Bea ☺
-------	-------
- S4: 

Ann ☺	Bea ☹
-------	-------

We have informally surveyed speakers about this example on a number of occasions. When asked to identify which situation at right accords with reality (or, the view of the passage's author), people reliably choose S4, in which just the names are reversed. This is Aloni's intended interpretation; the passage continues, "Bea is frowning on the right and Ann is smiling on the left." Why don't people choose S3, in which both the names and the facial expressions are reversed? This would seem, after all, to be more like "exactly the opposite" of Ralph's beliefs. The crucial factor is that the appositives *who is smiling* and *who is frowning*, despite being syntactically positioned below *Ralph believes*, are not interpreted there. To see what such interpretation would be like, we can look to simple conjunction, as in (5).

- (5) In front of Ralph stand two women. Ralph believes that the woman on the left is smiling and is Bea, and the woman on the right is frowning and is Ann. As a matter of fact, exactly the opposite is the case.

Here, syntactic position correlates with semantic scope; people reliably choose S3 as the one that characterizes reality, because all the conjuncts contribute semantically to the complement of the attitude predicate.

It is routine for expressives to scope outside of their embedded environments as well. Potts (2005) offers a number of attested cases. Here is a one involving an epithet, drawn from a weblog post; the author of this sentence is describing a complaint that he regards as frivolous:

- (6) The complaint says that the idiot filled in a box labeled "default CPC bid" but left blank the box labeled "content CPC bid (optional)".  
[[http://www.theregister.co.uk/2008/04/23/google\\_adfraud\\_court/](http://www.theregister.co.uk/2008/04/23/google_adfraud_court/)]

Naturally, the complaint does not contain the word *idiot*, as this would undermine its rhetorical position. In addition, the sentence preceding this in the weblog post says, "The plaintiff is an idiot who doesn't know what 'default' means", which presents the epithet's content as a direct predication. Thus, it is clear that *the idiot* encodes the speaker's dismissive opinion of the person who filed the complaint; whatever

content the expressive carries (Potts 2007; Constant et al. 2009), it is not part of the complement to *say* semantically. It is, rather, speaker-oriented.

On the basis of examples like these, Potts (2005) concludes that appositives and expressives are invariably speaker-oriented, regardless of syntactic configuration and discourse context. Wang et al. (2005) were, to our knowledge, the first to challenge this conclusion in print. Their short squib focuses on nominal appositives like *a professor, a famous one*, in which both the anchor *a professor* and the appositive clause *a famous one* are indefinite. Such appositives do indeed often give rise to non-speaker-oriented readings. For example, in (7), from a jaunty article about Alfred Kinsey, the biologist who founded the Institute for Research in Sex, Gender and Reproduction, the appositive *a Hoosier Dr. Mengele* is clearly intended to be interpreted inside the complement to *claim*.

- (7) Far out on the grassy knoll of sexology, there is a cult of procastity researchers who claim that the late Alfred Kinsey was a secret sex criminal, a Hoosier Dr. Mengele, who bent his numbers toward the bisexual and the bizarre in a grand conspiracy to queer the nation and usher in an era of free sex with kids. [20\_newsgroups corpus]

The example is rich in opposing emotive language: *grassy knoll of sexology* and *cult* contrast sharply with *Hoosier Dr. Mengele* and *the bisexual and the bizarre* to ensure that we have two distinct perspectives in play.

If the non-speaker-oriented cases were limited to doubly-indefinite examples like this, then we might be motivated to treat them as a special case. However, ARS find embedded readings with other kinds of appositive as well. In (8), for example, the appositive's anchor is a definite and the appositive is a full relative clause.

- (8) Joan is crazy. She's hallucinating that some geniuses in Silicon Valley have invented a new brain chip that's been installed in her left temporal lobe and permits her to speak any of a number of languages she's never studied. Joan believes that her chip, which was installed last month, has a twelve year guarantee.

Karttunen and Zaenen (2005) discuss related examples of embedding, which lead them to conclude, "we agree with Potts that supplemental expressions give rise to conventional implicatures but we disagree with his view that the author is always unconditionally committed to them." ARS reach a similar conclusion: "in the cases where CIs are anchored to an agent other than the speaker, they do appear to take narrow scope relative to the embedding attitude predicate [...]" (p. 738). In addition, they find non-speaker-oriented examples involving expressives, in environments similar to the one in (8):

- (9) [Context: We know that Bob loves to do yard work and is very proud of his lawn, but also that he has a son Monty who hates to do yard chores. So Bob could say (perhaps in response to his partner's suggestion that Monty be asked to mow the lawn while he is away on business):]

Well, in fact Monty said to me this very morning that he hates to mow the friggin lawn.

We conclude from these examples that Potts (2005) was wrong to claim that appositives and expressives are invariably speaker-oriented. In some utterances, they can describe the commitments of other individuals. The pressing question now is what conditions deliver such readings.

The literature offers two competing theoretical conceptions of how these non-speaker-oriented readings might arise:

- (10) a. **Configurational:** The source of non-speaker-oriented readings of appositives and expressives is semantic binding: their content can be bound by higher operators like attitude predicates, thereby shifting it away from the speaker (Schlenker 2003, 98, 2007, §4; Sauerland 2007).
- b. **Contextual:** The source of non-speaker-oriented readings of appositives and expressives is the interaction of a variety of pragmatic factors. In general, these interactions favor speaker-orientation, but other orientations are always in principle available, regardless of syntactic configuration (Potts 2007).

Hypothesis (10a) is a natural response to the above examples, and it has independent theoretical interest as well, since it would connect expressives and appositives with work on indexical shifting and logophoric pronouns (Schlenker 2003; von Stechow 2003; Anand and Nevins 2004; Oshima 2006). Hypothesis (10b) allows us to retain the multidimensional view of semantic composition developed in Potts 2005, but it calls for a more nuanced view of speaker-orientation and the pragmatics of appositives and expressives.

Potts (2007, to appear) is led away from (10a) and to (10b) by examples in which syntactically unembedded expressives receive non-speaker-oriented readings. Example (11) is a case in point. It is from an essay by Lewis Lapham, the liberal, populist essayist and political commentator who edited Harper's Magazine for many years. The sentence of interest is the last one. In it, he seems to call the beliefs of poor people "idiotic". This is out of step with what we know about Lapham, to say the least.

- (11) I was struck by the willingness of almost everybody in the room—the senators as eagerly as the witnesses—to exchange their civil liberties for an illusory state of perfect security. They seemed to think that democracy was just a fancy word for corporate capitalism, and that the society would be a lot better off if it stopped its futile and unremunerative dithering about constitutional rights. Why humor people, especially poor people, by listening to their idiotic theories of social justice?

[Lewis Lapham, Harper's Magazine, July 1995]

If we back up a few sentences, though, we see that a perspective shift has been achieved already going into this sentence. We have moved to the viewpoint of the senators and other powerful people involved. The adjective “idiotic” is meant to convey something about their views. There is no suggestion that they would overtly say or endorse such a characterization, but rather that their perspectives entail or justify it. The most important thing about this example, though, is that the adjective in question is not embedded at all. Whatever shift has happened is not one that is controlled by an attitude predicate.

Example (12) makes a similar point. Again, we need a lot of context to achieve the intended shift in the final sentence. Here, it is clear that the entire sentence is to be evaluated from a non-speaker perspective; the epistemic modal shifts along with the negatively charged *cronies*.

- (12) While shopping at one of my local Apple stores the other day, I overheard an earnest conversation about safeguarding Mac computers against things like viruses and trojans. The customer and companion were new to Mac life and were convinced that they should be very worried about viruses. The Apple salesperson on the floor repeatedly assured them that they would not need extra antivirus protection for their Mac. The customer then argued that Symantec makes an antivirus program for Macs, therefore, it must truly be a credible threat, otherwise there would be no such products. Some antivirus products are even sold in Apple stores. I’ve heard similar arguments before: if companies like Symantec or McAfee make antivirus applications for the Mac, then Macs must truly be vulnerable somehow, somewhere. Steve Jobs and the rest of the Apple cronies must be lying. [<http://news.digitaltrends.com/feature/79/antivirus-programs-for-mac-snake-oil-or-public-service>]

In light of these examples, we should return to ARS’s (8)–(9). Both use attitude predicates, but they also perform careful perspectival shifting in the lead-up to the target sentence. Examples (11)–(12) show that, at least for epithets, such contextual factors suffice to create non-speaker-oriented readings; while there is no doubt that attitude predicates facilitate shifting, they are not necessary for it to occur.<sup>3</sup>

We turn now to our experimental work, which bears directly on the choice presented by (10).

<sup>3</sup> The examples in this section bear a striking resemblance to Lasersohn’s (2005) examples in which a predicate of personal taste has shifted off of the speaker and onto another agent:

(i) [Context: “Suppose John is describing to Mary how their two-year-old son Bill enjoyed a recent trip to the amusement park. Something like the following dialog might occur:”]

Mary: How did Bill like the rides?

John: Well, the merry-go-round was fun, but the water slide was a little too scary.

We are reluctant to bring predicates of personal taste too close to expressives, though. We are persuaded by Lasersohn’s arguments that the two are different.

### 3 Experiment 1: Appositives and embedding

At the close of the previous section, we saw examples in which epithets receive non-speaker-oriented readings despite being in matrix clauses. We have not found comparable examples involving appositives, but our intuitions suggest that they are possible. Experiment 1 confirms these intuitions, thereby further supporting contextual hypothesis (10b) over configurational hypothesis (10a).

#### 3.1 Materials and method

The materials consisted of an invariant context and a target sentence containing an appositive clause (underlined in (13)). The experiment manipulated a single factor in the target sentence: whether the appositive was embedded under a verb of saying (A) or not (B). Our goal was to determine whether non-speaker-oriented readings arise outside of embedded contexts, and also to gauge whether embedding is a significant influence on such readings.

For all the successful non-speaker-oriented appositives and expressives we have seen, the speaker's perspective is sharply distinguished from that of another specific agent. We sought to reproduce this. All our contexts involve the speaker and a character with a notably distinct perspective from the speaker's. In the sample item (13), for example, the character is the roommate, and she is depicted by the speaker as paranoid.

(13) **Context:** I am increasingly worried about my roommate. She seems to be growing paranoid.

**Target:**

- A. The other day, she told me that we need to watch out for the mailman, a possible government spy.
- B. The other day, she refused to talk with the mailman, a possible government spy.

Whose view is it that the mailman might be a government spy?

**Response:**

- a. Mine (Speaker)
- b. My roommate's (Subject)
- c. Mine and my roommate's (Both)

Participants were instructed to select which agent's perspective was reflected in the appositive content. They were given three choices: speaker's (a), subject's (b), or speaker's and subject's (c) perspective. In (13), for example, they were asked whether the view that the mailman is a possible government spy should be attributed to me (the speaker), my roommate (the subject), or to me and my roommate (both). The presentation order of the responses was randomly generated in each question.

We sought to avoid calling too much attention to the target sentence in isolation from the preceding context. We think that the preceding context is a major factor in deciding the relevant judgment, so anything that encouraged subjects to judge the target sentence in isolation would likely detract from the pragmatic effect we are



testing. Thus, all the examples establish lexical and semantic connections between the target sentence's appositive and the probe question, but they don't explicitly focus attention away from the preceding context and onto the target.

The materials were divided into two balanced lists across four questionnaires, so that participants saw one and only one condition from each item. There were eight such pairs in all (given in appendix A), which were randomly interspersed with items from two other subexperiments (28 other experimental items, including 16 items from the experiment testing epithets described in Sect. 4) and four genuine filler items, for a total of 40 items per questionnaire. Of the eight experimental items, three pairs contained an appositive relative, and five pairs contained a nominal appositive.

## 3.2 Participants

Thirty-one undergraduate students and one graduate student from UMass Amherst participated in the study and were recruited in class or by email from the authors. Undergraduates received course credit for their participation. The study was conducted over the internet, and participants were asked to find a quiet location to complete the questionnaire. All the participants self-identified as native speakers of English.

## 3.3 Data analysis

### 3.3.1 *Distribution of responses*

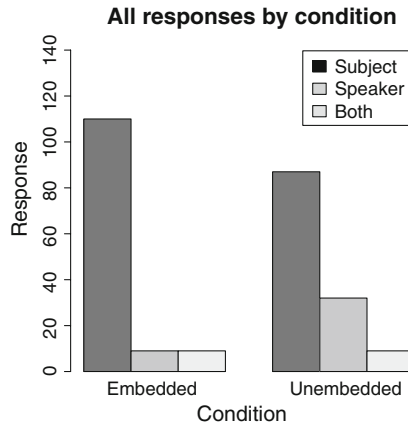
Participants strongly preferred non-speaker-oriented interpretations of appositive content in both conditions: the Subject interpretation was selected 86% of the time for Embedded conditions and 68% of the time for Unembedded conditions. Speaker interpretations were more prominent in Unembedded conditions (25%) than Embedded conditions (7%). The Both response was selected 7% of the time in both the Embedded and Unembedded conditions. Importantly, the Subject response was vastly preferred regardless of condition, and was selected as the response in 77% of the total number of trials. Figure 1 summarizes the distribution of the responses.

### 3.3.2 *The influence of syntactic position*

Non-speaker-oriented readings were strongly favored in both the Embedded and Unembedded conditions, but they arose more often in Embedded conditions. Is this difference significant? That is, does syntactic embedding significantly influence whether people choose non-speaker-oriented readings? To answer this question, we subjected the data to a linear mixed-effect logistic regression analysis, treating Participants and Items as random effects.<sup>4</sup> (For discussion of logistic regression in the context of analyzing psycholinguistic data, see Baayen 2008; Jaeger 2008.) All

---

<sup>4</sup> A natural alternative to use here is the chi-squared test, but the mixed effects model allows us to bring Item and Participant variability into the statistical analysis, whereas the chi-squared test ignores it.



(a) Total number of response scores

Condition	Response		
	Subject	Speaker	Both
Embedded	110 (86%)	9 (7%)	9 (7%)
Unembedded	87 (68%)	32 (25%)	9 (7%)
Combined conditions	197 (77%)	41 (16%)	18 (7%)

(b) Proportion of scores

**Fig. 1** Experiment 1 response data, by condition

analyses were conducted with the R statistical computing software (R Development Core Team 2008).

For this analysis, we grouped the Both and Speaker responses into a single category of Non-Subject readings, as in Table 1. We took this step because of the indeterminacy inherent in the Both responses. While it is possible that, in these cases, the content is genuinely perceived as a commitment of both speaker and subject, we suspect that these readings are the result of pragmatic enrichment: either the experimental subject attributed the content in question to the Subject and then inferred that the Speaker believed it also, or the reverse (see Karttunen's 1973 discussion of how presupposition plugs can 'leak'). Thus, we are unsure of how to count Both responses when the focus is on Subject readings, so we have decided to take the conservative approach of treating them as Non-Subject readings, even though we might lose genuinely Subject-oriented readings in this conflation.<sup>5,6</sup>

<sup>5</sup> It might be argued that the Both responses represent uninterpretable noise in the data. To address this concern, we fitted another linear mixed-effects logit model with Both responses excluded. Excluding the scores did not influence the overall results: Subject interpretations of the appositive were still significantly more likely when the appositive was embedded;  $p < 0.001$ .

<sup>6</sup> An anonymous reviewer observed that item 5 from our materials was slightly different from the others in that it does not explicitly introduce the speaker into the context. However, removing the item did not affect the analysis.

**Table 1** The experiment 1 response data with the Speaker and Both categories combined into a single Non-Subject category

	Response	
	Subject	Non-Subject
<i>Condition</i>		
Embedded	110 (86%)	18 (14%)
Unembedded	87 (68%)	41 (32%)
Combined conditions	197 (77%)	59 (33%)

**Table 2** Experiment 1: linear mixed-effects logit model

	Estimate	Std. Error	Wald Z	<i>p</i> value
(Intercept)	1.14	0.58	2.00	<0.05
Embedded	1.44	0.36	3.90	<0.001

We first report the model that addresses the central manipulation, i.e., whether embedding of an appositive influenced its interpretation. The results, displayed in Table 2, show that participants were indeed significantly more likely to interpret embedded appositives as Subject oriented (86% of trials) than they were unembedded appositives (68% of trials);  $z = 3.90, p < 0.001$ . We can therefore conclude that (at least for these examples) syntactic embedding positively correlates with non-speaker-oriented readings.

We now describe the model in Table 2 by interpreting its coefficients. The model provides an estimate of the coefficient parameter for each fixed effect. The Intercept parameter in this model estimates the probability of a Subject interpretation when the appositive is *not* embedded, modulo influence from the random effects, which capture variance due to differences between individual participants and items in our model. The Embedded parameter estimates the effect on the probability of a Subject interpretation in log odds when the parameter is set to 1, i.e., when the appositive is embedded, again modulo random effects. Thus, the model predicts that, within the sample we tested, a Subject interpretation of the appositive was 1.44 log odds more likely when embedded. Transforming log odds to probabilities, embedded appositives were 80% more likely to be interpreted as Subject-oriented in comparison to unembedded counterparts.<sup>7</sup>

The Wald Z value, which is obtained by dividing the parameter estimate by the standard error, estimates how far each coefficient is from zero, normalized in terms of its standard error (Wald 1943; Jaeger 2008). The rightmost column depicts the probability that a Z value of this size would be obtained by chance. Thus, embedding was shown to positively contribute to the likelihood of a Subject interpretation of an appositive.

<sup>7</sup> We transform the log odds of some variable  $v$  into probability space by  $P(v) = \text{logit}^{-1}(v) = \frac{1}{1+e^{-v}}$ .

**Table 3** Experiment 1: modeling the effect of the presence of a PPT in prior context with embedding

	Context	
	No PPT	PPT
Unembedded	15 (31%)	72 (90%)
Embedded	39 (81%)	71 (89%)

(a) The distribution of Subject-oriented interpretations from the questionnaire, organized by Embedded and PPT factors.

	Estimate	Std. Error	Wald Z	<i>p</i> value
(Intercept)	-0.84	0.50	-1.67	0.09
Embedded	2.39	0.50	4.77	<0.001
PPT	3.26	0.71	4.62	<0.001
Embedded:PPT	-2.53	0.73	-3.46	<0.001

(b) This linear mixed-effects logistic regression model tests whether embedding an appositive and the presence of a PPT in prior context significantly correlates with the proportion of Subject-oriented responses. Both factors significantly contributed to an increased likelihood of Subject-oriented responses. Random effects are omitted for convenience.

Is embedding the *only* factor to reliably contribute to Subject-orientation? In Harris and Potts (2009), we conducted a post-hoc analysis of the data and identified the context description as an important factor: where the description contains a predicate of personal taste, Subject judgments are more frequent.<sup>8</sup> It turns out that a model with an interaction term between Embedded and PPT (1 if the context contains a predicate of personal taste, else 0) and a single Item random effects term provides the best model, i.e., the best balance between goodness of fit and model complexity.<sup>9</sup> The distribution of Subject-oriented interpretations partitioned along these lines is provided in Table 3a.

We again fit a logistic linear mixed-effects logit regression to the data with PPT and the interaction term Embedded:PPT as additional factors and Item as the only random effects term, as shown in Table 3b. We observed both a main effect of embedding,  $z = 4.77$ ,  $p < 0.001$ , and a main effect of the presence of a PPT,  $z = 4.62$ ,  $p < 0.001$ , as well as an interaction between these two factors,  $z = -3.46$ ,  $p < 0.001$ .

For illustration, we show that the parameter estimates from this model tightly correspond to the response distribution in Table 3a. We use logistic regression to predict the probability of Subject readings in terms of the syntactic position of the appositive ( $x = 1$  if Embedded, else 0), the presence of PPT in the preceding context ( $y = 1$  if PPT, else 0), and their interaction ( $z = 1$  if  $x = y = 1$ , else 0). The fitted model of fixed effects terms is provided in (14); see also Footnote 7.

<sup>8</sup> Predicates of personal taste, including *paranoid*, *absurdly optimistic*, *huge snob*, *crazier than ever* and *outlandish*, were included in five of the eight items, and were not orthogonal to the central manipulation.

<sup>9</sup> The models reported here were first compared to more complex models, using the top-down model-fitting technique discussed in Crawley (2007). We retained the simplest model that did not significantly reduce the quantity of variance explained.

$$(14) \quad \text{Pr}(\text{Subject}) = \text{logit}^{-1}(-0.84 + 2.39x + 3.26y + -2.53z)$$

This model says that the probability of a Subject-oriented interpretation of an embedded appositive ( $x = 1$ ), without a predicate of personal taste in the context ( $y = z = 0$ ), is  $\text{logit}^{-1}(-0.84 + 2.39) = \text{logit}^{-1}(1.55) = .82$ , while an unembedded appositive ( $x = 0$ ) in the same context is  $\text{logit}^{-1}(-0.84) = .30$ . The predicted values correspond nearly exactly to the observed values, 81% and 31%, respectively, from Table 3a. Similarly, there is a close correspondence between predicted and observed values for the PPT factor: the model predicts the likelihood of a Subject-oriented interpretation for unembedded appositives ( $x = z = 0$ ) with a PPT in the prior context ( $y = 1$ ) to be  $\text{logit}^{-1}(-0.84 + 3.26) = \text{logit}^{-1}(2.42) = .92$ , and embedded appositives in a similar context ( $x = y = z = 1$ ) to be  $\text{logit}^{-1}(-0.84 + 2.39 + 3.26 + -2.53) = \text{logit}^{-1}(2.28) = .91$ .

We believe that describing the above model in this way is informative for a variety of reasons. First, it underscores that other factors besides embedding contribute to the perspectival orientation of appositive clauses. Second, it suggests that multiple factors, such as embedding and perspectivally-charged contexts, can simultaneously contribute to the interpretation of subjective content, especially when contravening conventional usage (e.g., away from speaker-orientation in the case of appositives). Third, as evidenced by the significant interaction between Embedded and PPT parameters, it raises the possibility that such factors may combine to strengthen this non-conventional interpretation, recalling Smith's (2009) observation that subjective terms 'additively' give rise to subjective interpretations that shift from the Speaker to another character in the text.

In order to avoid over-interpreting the data, we have not fitted additional models. However, we emphasize that the factors we have discussed may not exhaust those contributing to perspectival orientation. We invite others to examine the data for themselves and, for that purpose, make the data available upon request.

### 3.4 Discussion

Experiment 1 supports four central conclusions:

- (15) a. Non-speaker-oriented readings are available for appositives, both when they are syntactically embedded inside attitude predications and when they are in matrix clauses.
- b. In perspectivally-rich contexts, non-speaker-oriented readings are even preferred under some circumstances.
- c. Embedding inside an attitude context significantly increases the likelihood of a non-speaker-oriented reading.
- d. Other factors besides embedding contribute to the likelihood of a non-speaker oriented reading, and may interact with embedding or other more contextual cues.

These conclusions challenge the view of Potts (2005) that appositives are invariably speaker-oriented. The experimental items involve both nominal

appositives and appositive relatives, with a variety of different morphological forms, positioned in various places in the main clause. We expect that these findings reflect the general state of affair for these clauses.

Conclusions (15a) and (15d) support the contextual hypothesis (10b) and challenge the configurational hypothesis (10a), which ties non-speaker-oriented readings too closely to the presence of an embedding operator. Our findings are consistent with a theoretical position on which we enrich Potts's (2005) multidimensional view of semantic composition with a theory of perspective taking in context (Gianakidou and Stavrou 2008; Lasersohn 2005). However, hypothesis (10a) is partly vindicated: it seems that embedding is a significant positive influence on non-speaker-oriented readings.

In (15b), we see a glimpse of which factors guide speakers to interpret a given appositive as non-speaker-oriented, but we still do not have a clear picture of the details. However, we have made progress on this question for the case of epithets. In the next experiment, we examine how contextual factors can influence perspectival orientation for such expressives.

## 4 Experiment 2: Epithets and perspective shift

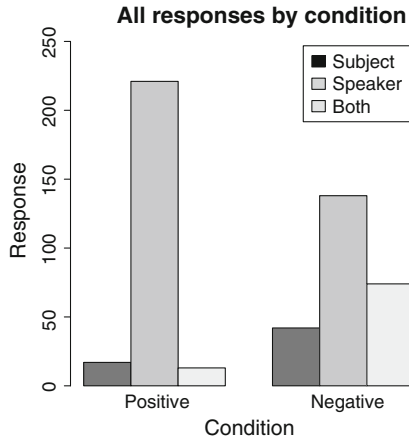
At the end of Sect. 2, we saw that non-speaker-oriented readings are available even for unembedded epithets like *the jerk*. Harris (2009) reports on a pilot study that further supports this basic finding and also begins to identify the pragmatic factors that favor one orientation over the other. Experiment 2 follows up on this pilot study with a more complex manipulation.

### 4.1 Materials and method

This experiment tested whether the emotional content associated with epithets could be non-speaker-oriented in unembedded environments.<sup>10</sup> After Harris (2009), we reasoned as follows: if participants were given evidence that the subject of the attitude report held a negative emotive stance towards the referent of an epithet, they would more often interpret that epithet as non-speaker-oriented. Conversely, if they were not given such biasing evidence, then they would more often favor speaker-orientation, which is arguably the default strategy.

We further hypothesized that the stronger the evidence for an emotional relationship between the attitude holder and referent of the epithet, the more likely participants would be to perceive a non-speaker-oriented reading. We sought to strengthen this evidence by using intensives like *really*, *totally*, and *super* (Beaver and Clark 2008; Potts and Schwarz 2008) to amplify the modifiers used to create the bias; it is one thing to be upset (or delighted) that *X is fantastic* and another to be upset (or delighted) that *X is totally fantastic*.

<sup>10</sup> While epithets do not convey an *exclusively* negative emotive relation between the referent and the attitude holder, we did not detect any interpretations in our items in which a positive emotional stance was supported in the context provided. We assume that our participants would not entertain any such interpretations, as well.



(a) Total number of response scores

Condition	Response		
	Subject	Speaker	Both
Positive	17 (7%)	221 (88%)	13 (5%)
Negative	42 (17%)	138 (54%)	74 (29%)
Combining conditions	59 (11%)	359 (71%)	87 (17%)

(b) Proportion of scores

Fig. 2 Experiment 2 response data, by condition

Thus, the experimental design crossed two factors: the polarity of Context (Negative, Positive) with the presence of an Intensifier (Y, N). The central hypothesis mainly concerns the effect of Context on the availability of non-speaker-oriented interpretations. We did not, for instance, predict that these two factors would interact in a meaningful way. In particular, we did not expect the Intensifier manipulation to affect either Positive or Negative contexts differently from the other, although we regarded this result as very possible.

The experimental items consisted of sixteen quadruplets, with the same pattern as the sample item (16). Each varied only in whether the context was positive or negative and whether there was an intensifier before the adjective or not. The materials were presented in the same experiment as the materials from experiment 1, described in Sect. 3. Appendix B gives the full set of non-filler materials.

(16) **Suppose you and I are talking and I say:**

- A. My classmate Sheila said that her history professor gave her a low grade. (Negative, N)
- B. My classmate Sheila said that her history professor gave her a really low grade. (Negative, Y)
- C. My classmate Sheila said that her history professor gave her a high grade. (Positive, N)

- D. My classmate Sheila said that her history professor gave her a really high grade. (Positive, Y)

**Target:** The jerk always favors long papers.  
Whose view is it that the professor is a jerk?

**Response:**

- a. Mine (Speaker)  
b. Sheila's (Subject)  
c. Mine and Sheila's (Both)

With our materials, we sought to balance a number of pressures. First, as in experiment 1, we needed the target sentence to be judged in its discourse context, rather than in isolation. Thus, the target and the question are linked lexically and semantically, but we did not otherwise call attention to the target in isolation from what preceded it. Second, expressives like *the jerk* can be hard to use neutrally, even when they are quoted, so there was a danger that our questions would involve new expressive acts and thus introduce new perspectival complications. We tried to manage this by using indefinite forms of them (e.g., *a jerk* rather than *the jerk*) and querying that content. Third, epithets tend to be highly referentially ambiguous. We always intended our epithets to refer to the subject of the embedded clause in the preceding sentence, but some speakers might have perceived readings where the epithet referred to the subject of the matrix clause of the preceding sentence. Such readings basically ensure speaker-orientation and thus have the potential to diminish the effect we were testing for. Both the contexts and the following questions were designed to disfavor such readings as much as possible, though they might still be favored for information-structuring reasons.

## 4.2 Participants

The study involved the same thirty-two participants from experiment 1, as described in Sect. 3.2

## 4.3 Data analysis

### 4.3.1 The distribution of responses

We first examine our data by the Context condition. Unlike in experiment 1 on appositives, speaker-oriented interpretations were preferred across conditions. However, while 88% of the judgments for the Positive condition were speaker-oriented, just 55% were speaker-oriented in the Negative condition. Conversely, subject-oriented responses increased to 17% in the Negative condition, up from 7% in the Positive condition. Similarly, Both responses rose to 29% in the Negative condition, up from 5% in the Positive condition. This response data is summarized in Fig. 2.<sup>11</sup>

<sup>11</sup> A typo affecting the first nine participants in one item was found. Responses for this item were coded as "NA" and removed from further analysis.



**Table 4** Proportion of scores for experiment 2, crossing Context by Intensifier

Condition	Intensifier	Response		
		Subject	Speaker	Both
Positive	N	8 (3%)	114 (45%)	4 (1%)
Positive	Y	9 (4%)	107 (43%)	9 (4%)
Negative	N	17 (7%)	73 (29%)	38 (15%)
Negative	Y	25 (10%)	65 (25%)	36 (14%)

Table 4 gives the overall distribution of responses with the Intensifier condition included. It is clear from the proportions that Intensifier did not successfully manipulate the availability of Subject responses, as confirmed in the analysis below. As a result, we do not consider the Intensifier condition in much detail.

#### 4.3.2 The influence of context

As we noted above in connection with Table 4, the Intensifier variable was not a significant predictor of how an epithet was interpreted. This was confirmed by fitting a linear mixed effect logit model to the data, again treating Participants and Items as random effects. Neither the Intensifier manipulation nor its interaction with the Context condition had a significant coefficient. This is somewhat evident in Table 4 for Speaker responses as well: rather than enhancing the bias created by the Context, the Intensifier seemed to diminish it slightly, though we don't have evidence that this is reliable. For this reason, the Intensifier manipulation has been removed from the statistical models.

We again face a choice about how to view the Both data. As in experiment 1, we opt to combine the Speaker and Both responses into a single category of Non-Subject responses. Treating the scores this way biases against our hypothesis, because it reduces the difference between Positive and Negative Contexts, as shown in Table 5a.

The results of the second questionnaire confirm the pattern reported in Harris (2009): contexts with negative predicates modifying the antecedent of the epithet were significantly more likely (17%) to be given a Subject response than relevant counterpart items (7%) in the experiment;  $z = 3.69$ ,  $p < 0.001$ .

#### 4.4 Discussion

The results of experiment 2 further support the claim that non-speaker-oriented readings are possible for expressives, if the right contextual factors are present. The results also suggest that such readings do not require syntactic embedding, and thus they further challenge the configurational hypothesis (10a).

The experiment manipulated carefully controlled contexts so that changing a single word, and thereby alternating the general positive or negative emotive polarity of the context, triggered a series of inferences about whether the subject of

**Table 5** Experiment 2: the organization of the response data for the questionnaire logit analysis, with details of the logit model

	Context	
	Pos	Neg
Subject	17 (7%)	42 (17%)
Non-Subject	234 (93%)	212 (83%)

(a) The questionnaire response data with the Speaker and Both categories combined into a single Non-Subject category.

	Estimate	Std. Error	Wald Z	<i>p</i> value
(Intercept)	-4.02	0.54	-7.45	< 0.001
Neg	1.32	0.36	3.69	< 0.001

(b) This linear mixed-effects logistic regression model tests whether the nature of the context (Positive vs. Negative) significantly correlates with the proportion of Subject-oriented responses. The Neg condition significantly affected the likelihood of increased Subject-oriented responses.

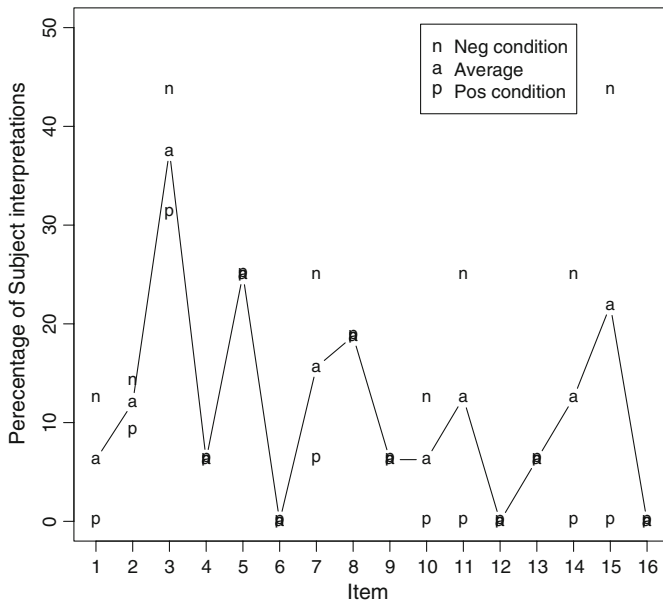
the attitude report stood in a negative relation to the referent of the epithet. The response data suggest that these inferences affected how speakers understood the epithets. (We did not, though, find evidence that intensives reliably contribute here.)

A reviewer points out that there is intuitive variation in the experiment 2 items that could impact perspectival orientation. For example, some of our scenarios depict close personal relationships between the participants, whereas others involve more distant ones. There was also probably variation in how familiar participants were with the epithets; to avoid offending our subjects, we stayed clear of extremely caustic, but widely used, epithets like *the asshole*, which led us to less familiar ones like *the skinflint*. It is therefore worth looking at our results by item. Figure 3 summarizes. The items are listed along the *x*-axis, with numbers corresponding to the listing in Appendix B. The *y*-axis lists the percentage of Subject interpretations, with ‘n’ marking negative responses, ‘p’ marking positive responses, and ‘a’ giving the average of the two. Only items 1, 2, 3, 7, 10, 11, 14, and 15 show the effect described in Sect. 4.3.2, though we have no evidence that any items go against our hypothesis. In fact, some items not showing the effect nonetheless received high percentages of Subject responses. The bottom line is that our hypothesis shines forth despite this variation in our materials, so there is no need to exclude any items. We do think, though, that this information could prove useful for designing future studies on the factors that govern perspectival orientation.

## 5 Corpus study

The results of experiment 1 (Sect. 3) suggest that embedding attitude predicates, while not required for non-speaker-oriented readings of appositives, can facilitate such readings. The present section shifts the emphasis slightly. Now that we know speaker-oriented and non-speaker-oriented readings are both possible, we would like

**Percentage of Subject responses, with deviation by condition**



**Fig. 3** Experiment 2 by-items analysis. The item numbers correspond to those in Appendix B

to know what people are actually doing with their appositives. In particular, how frequent are non-speaker-oriented readings in naturally occurring text? We address this question with a new corpus of embedded appositives (Potts and Harris 2009).<sup>12</sup>

### 5.1 Data sources

This data set is concerned only with appositive relatives with nominal anchors that appear in the syntactic complements to attitude verbs and verbs of saying, as in (17).

- (17) Hartzenberg said he would ask Terre'Blanche, who heads the extremist Afrikaner Resistance Movement (AWB), if he would meet Mandela.

The data are drawn from the following sources, all of which are freely available on the Net:

- (18) a. The Gigaword fragment downloadable from the LDC website
- b. The 20\_newgroups corpus
- c. The sample of the Penn Treebank distributed with NLTK
- d. Novels downloadable from Project Gutenberg
- e. TV show transcripts downloadable from CNN.com

<sup>12</sup> The corpus distribution includes all the data in XML format as well as a browser-based search function, our own annotation manual, the instructions that the assessors received, and screenshots of the browser-based annotators used by us and by the assessors.

In all, this amounts to about 177 million words in a variety of styles and genres. The first step in unearthing the relevant examples was pulling out all the sentences that match the following regular expression:

(19) `PREDS(es | s | ed)? . * [A-Z] [a-z]*\ , \ s + (who | whom | whose) \ s`

Here, PREDS stands in for a long disjunctive list of attitude predicates:

(20)

affirm	allege	announce	argue
assert	believe	certain	claim
conclude	conjecture	declare	guess
imply implies	judge	move	presume
request	remark	report	rumor
say said	suggest	suppose	sure
think thought	trust	understand	ask
question	request	query queries	

We then went through these results by hand, throwing out spurious matches. This left us with 278 examples.<sup>13</sup>

## 5.2 Annotation

With the examples collected, we conducted a two-stage annotation project. In stage 1, we went through the examples ourselves (first independently, and then together), looking for *textual* evidence for the intended interpretation of the appositive and summarizing that evidence in short prose statements. In stage 2, our evidence was assessed by two independent annotators. This process is described in more detail in the next two subsections.

### 5.2.1 Stage 1: Seeking textual arguments

In stage 1, we were the annotators. We had two related tasks. The first was to answer the question ‘At which level does the speaker intend the appositive to be interpreted?’ The choices for this annotation were as follows:

- (21)
- a. Textual evidence for text-level
  - b. Textual evidence for embedded
  - c. No textual evidence detected; seems text-level
  - d. No textual evidence detected; seems embedded
  - e. No textual evidence detected; intentions unclear

We deliberately focused the annotation task on speaker intentions. One reasonable conclusion to draw from experiment 1 is that all embedded appositives are

<sup>13</sup> At this stage, we also tagged the relevant examples for the root form of the embedding verb.

semantically underspecified for the perspective of the appositive, so the interesting question is probably not what is allowed, but rather what was actually intended (or what an audience is likely to assume the speaker intended).

Where we judged the example to have textual evidence (the first two options in (21)), we summarized that evidence succinctly, citing the relevant passages and seeking to explain their relevance to the question at hand. Here are two examples, the first arguing for a text-level interpretation, the second for an embedded interpretation; these examples are available along with their full textual contexts as part of the corpus distribution.

- (22) a. **Target sentence:** A government prosecutor said Wednesday he plans to drop vandalism charges against a Malaysian teenager allegedly involved in a spate of spray painting cars with a young American, Michael Fay, who was caned recently.
- b. **Evidence:** A later sentence elaborates on the details of Fay’s punishment: “[...] was given four strokes of a rattan cane two weeks ago [...]”.
- (23) a. **Target sentence:** Israel says Arad was captured by Dirani, who may have then sold him to Iran.
- b. **Evidence:** The connective ‘then’ presupposes that if the complement clause is interpreted as embedded, then so too must the appositive. There is no evidence contraindicating an embedded reading of the complement clause. The epistemic modal also seems to be anchored to the subject, Israel.

We also included in the data some arguments that we considered to be poor, either because they did not restrict attention to textual evidence or because they did not necessarily support a particular judgment. We did this as a cautionary measure; in the event that our textual arguments were all good ones, we wanted to prevent our assessors from growing accustomed to providing positive assessments.

In general, finding and formulating textual arguments was demanding. It required a close-reading of the text as well as sensitivity to a wide variety of textual clues. Very often, the author’s intentions were clear but finding evidence for those intentions was considerably more challenging. We did this work ourselves, rather than passing it along to our annotators, because we wanted to decouple the task of finding textual evidence from the task of assessing that evidence and using it to make a judgment about the author’s intentions. We were the detectives, and we let our assessors judge what we found.

Though it is not sensible to give a comprehensive list of all the different kinds of evidence we used, we can outline some of the general considerations that we had in mind as we worked:

- (24) a. Presupposition triggers dependent on the truth of the appositive for definedness (schematically: *A begin P* in the appositive and *A stop P* later in the text, at the matrix level).

- b. Text-level anaphoric devices (especially sentential *it*, *that*, and *this*) that depend anaphorically on the appositive.
- c. Patterns of contradiction in the broad sense defined by de Marneffe et al. (2008) (some of which involve presupposition triggers).
- d. Discussion that pragmatically presupposes the truth of the appositive content.
- e. Independent text-level assertions of the appositive content (fairly common in newspaper articles, which summarize at the start and then expand on that summary throughout the remaining text).

We emphasize *textual* in all this. In many of the cases where the given annotation is ‘no textual evidence detected’, the content seemed clearly intended to be interpreted as text-level. Such content tends to be a matter of historical record, or was intended to be given as a piece of new, albeit secondary content. Here are two examples:

- (25) a. The king said that he and his wife were “greatly saddened” by the death of Onassis, widow of former US president John Kennedy, who was assassinated in Dallas in November 1963.
- b. Texas Air Corp said it named Norman McInnis as president of its Britt Airways unit, succeeding Bill Britt, who retired March one.

In neither of these cases did the text verify the appositive’s content independently, so they were included in the ‘no textual evidence’ category.<sup>14</sup>

Our goal in relying on textual evidence was to avoid circularity of argumentation. If the evidence for text-level interpretation is merely that our intuitions point to such an interpretation, then the judgment is easily challenged and hard to further support. Relying on textual evidence means that we have somewhat more objective criteria for the annotations, though it should be kept in mind that these criteria are intricate and depend on defeasible inferences. For highly relevant discussion, see Zaenen et al. (2005) and Manning (2006).

### 5.2.2 Stage 2: Independent assessment

Our assessors worked with a browser-based annotation tool that we designed. It presented the appositive in its textual context, highlighted in yellow. It also presented our textual evidence. The assessor’s job was to answer two questions:

- (26) What kind of reading is this evidence for?
- a. Text
  - b. Embedded
  - c. Unclear

<sup>14</sup> The corpus contains many examples like (25b), in which all the central facts are reported under verbs of saying, which effectively act as evidential markers (Bergler 1992; Simons 2007).

- (27) What is the status of the textual argument?
- a. Good
  - b. Bad

There was also an optional text box in which our assessors could comment on the argument. They were encouraged to do so. An example of these comments may be found in Appendix C.

### 5.3 Analysis

Appendix C contains a full example from the corpus, to provide a better sense for what the examples and annotations are like. The entire corpus is available on the Internet (Potts and Harris 2009).

There are a number of ways in which we can view the annotated data. Our primary goal for this section is to use the data to assess how frequent non-speaker-oriented readings are in real discourse. We thus focus on the most restrictive subset of the annotated data: the set of examples for which we have three-way agreement (Harris and Potts, plus our two assessors) that the textual evidence found for them was rated ‘Good’. There are 34 such examples in our data, just two of which are non-speaker-oriented (Embedded, in the annotation scheme): example (7) and example (23). The remaining 32 examples are speaker-oriented (Text). We cannot think of any reasons why the demand for text-level interpretations would bias in favor of non-speaker-oriented interpretations. Indeed, if anything, we would expect non-speaker-oriented interpretations to be more likely to appear with textual evidence, since the speaker is likely to feel some pressure to disavow the content. And yet still the bias is heavily in the other direction, with just 6% of the examples plausibly intended as non-speaker-oriented.

Before we close, a few words on inter-annotator (inter-assessor) agreement are in order, as these provide some sense for how reliable the annotations are. We calculated two different kappa scores of agreement (Cohen 1960). For both these calculations, we left ourselves out of the picture, focusing instead on the level of agreement between the two assessors.

If we consider our two assessors as selecting, for each example, a single annotation from the set {‘Good’, ‘Bad’} and another single annotation from the set {‘Text’, ‘Embedded’, ‘Unclear’}, then there are six possible categories. The kappa measure for our assessors is, from this perspective, about 39%; Table 4 provides the details. This is a moderate level of agreement, not especially encouraging on its own, but arguably reasonable given the difficulty of the task.

If we restrict attention to the examples for which both assessors regarded the status of the argument as Good, then we can measure the degree to which they agreed on what the evidence was good for: Embedded, Text, or Unclear. This is the upper left quadrant of Table 6. Here, we have 100% agreement. Taken together, these kappa results suggest that there is uncertainty about which arguments are good and which are bad, but that where there is agreement that the argument is good, the judgments about the nature of that evidence are highly reliable.

Though we know of no annotation efforts that are directly comparable to this one, two recent projects are worth briefly describing here for the purposes of informal

**Table 6** Inter-assessor agreement

		Good			Bad		
		Text	Embedded	Unclear	Text	Embedded	Unclear
Good	Text	32	0	0	11	0	2
	Embedded	0	5	0	0	2	0
	Unclear	0	0	0	0	0	0
Bad	Text	10	1	0	11	0	1
	Embedded	0	3	0	1	4	0
	Unclear	0	0	1	1	1	0

The kappa measure for the entire table is a modest 39%, but if we restrict attention to the examples for which both assessors thought our evidence was good (upper left quadrant), then the agreement rate is 100%

comparison. First, de Marneffe et al. (2008) annotated a range of different texts for a broad construal of contradiction, embracing not only logical inconsistencies but also pragmatic clashes arising from implicit speaker meaning. This is a very different annotation task than ours, but the two efforts share three properties: (i) the corpora were diverse, (ii) the annotators were confronted with very general problems in textual understanding, and (iii) the phenomena in question are relatively infrequent. de Marneffe, Rafferty and Manning report a kappa score of 81%, which is notably higher than what we achieve for our full data set, though in-line with our result when we focus on the ‘good argument’ cases. Second, the FactBank Corpus (Saurí and Pustejovsky 2009) extends the Penn Discourse TreeBank Corpus (Miltsakaki et al. 2004) with a wide range of additional semantic annotations, including information about speaker commitments for matrix and embedded clauses. For the annotations concerning speaker commitments, Saurí and Pustejovsky report kappa scores of 81%. However, their task strikes us as substantially easier than our own in large part because the annotators could rely heavily on lexical information. Indeed, the FactBank annotation manual is dominated by instructions about how to handle specific lexical cases. Interestingly, though, they report that attitude predicates like *believe*, *think*, and *decide* remained a prominent source of errors in virtue of the complex ways in which they convey perspectival and evidential information.

We hope that, in making our corpus, annotations, manuals, and instructions available, we are helping to pave the way to higher-fidelity annotations in the future. For additional discussion of the challenges facing tasks like this (and a criticism of existing data sets with similar ambitions), we refer to Zaenen et al. (2005).

## 5.4 Discussion

A speaker who utters an appositive with the intention of having it be understood as non-speaker-oriented has undertaken a risky communicative strategy in the following sense: it runs counter to hearer expectations about how these constructions will be used. Thus, this is a reliable strategy only in contexts that are rich enough to



support another perspective in just the right ways. To put it another way: you might always be free to intend your appositive to be understood as non-speaker-oriented, but your audience will often be unable to recover your intentions. And you probably know that your addressee will be unable to recover your intentions, which will lead you to adopt another strategy (assuming you wish to be understood). So, in this broadly game-theoretic fashion, we arrive at the apparent defaults evident in our corpus results, while at the same time allowing that contextual factors like those manipulated in our experiments might lead to different outcomes.

We think this imbalance has an underlying theoretical cause. Suppose appositives and expressives are inherently underspecified for their orientation. Semantically, we might achieve this by including a free variable that determines the *epistemic anchor* or *judge* for the content in question (Farkas 1992; Giannakidou 1999; Giannakidou and Stavrou 2008; Lasersohn 2005). Since there is no general morphological convention for specifying this information directly, it must always be left to the context. Appositives have many of the morphosyntactic and intonational properties of regular asserted declaratives, which are also overwhelmingly speaker-oriented, so perhaps it is unsurprising that appositives are generally speaker-oriented as well.

## 6 Conclusions

Potts (2005, 1) writes, “I hope readers of this book are struck by how little pragmatics it contains”. ARS take him to task for this, arguing that the important questions about appositives and expressives are largely pragmatic. At this point, we are inclined to agree; a unifying theme of the experimental and corpus work described here is that the important, challenging interpretive questions about appositives and expressives concern where and how they are used.

Experiment 1 shows that non-speaker-oriented readings of appositives arise even outside of embedded contexts, thereby calling into question any account of such readings that depends on semantic mechanisms. Experiment 2 enhances this general conclusion by identifying a significant contextual predictor of the interpretive orientation of epithets like *the idiot*. Taken together, these findings suggest that we should look to the discourse, rather than to the logical forms, to determine how these constructions are understood. Whatever indexicality inheres in these phenomena seems to more closely resemble the discourse-based logophoricity of Kuno (1987) and Pollard and Sag (1992) (see also Buring 2005, Chap. 11) than the bound indexicals of Schlenker (2003), von Stechow (2003), Anand and Nevins (2004), and related work.

The corpus work complements the experimental work, by suggesting that speaker-oriented readings of appositives dominate in real discourse, even when the appositive is syntactically positioned in such a way as to potentially favor non-speaker-orientation. The bias looks significant enough that we should seek a theoretical understanding of it, perhaps by looking away from categorical generalizations and towards those that are based in speaker and hearer expectations and the relative pragmatic stability that they create (Lewis 1969, 1975).

**Acknowledgments** For valuable discussion and advice, we thank Patricia Amaral, David Beaver, Charles Clifton, Jr., Lyn Frazier, Craig Roberts, Philippe Schlenker, Florian Schwarz, E. Allyn Smith,

Adrian Staub, and audiences at OSU, Stanford, and UMass Amherst. Thanks also to Chris Davis and Aynat Rubinstein for assessing our corpus evidence. This material is based upon work supported, in part, by the National Science Foundation under Grant No. BCS-0642752. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

## Appendices

### A Appositive materials

The experiment consisted of just one factor, which manipulated whether the appositive was embedded in an attitude report or not. Target sentences for each condition were preceded by the same sentence, which provided some context for the target. The slash (/) separates Embedded conditions from Unembedded conditions.

1. I am increasingly worried about my roommate. She seems to be growing paranoid. The other day, she told me that we need to watch out for the mailman, a possible government spy. / The other day, she refused to talk with the mailman, a possible government spy.
2. My friend Sal is absurdly optimistic. He told me that the lottery ticket he bought yesterday, a sure winner, is the key to his financial independence. / All he could talk about at dinner was the lottery ticket he bought yesterday, a sure winner.
3. My aunt is extremely skeptical of doctors in general. She says that dentists, who are only in it for the money anyway, are not to be trusted at all. / Dentists, who are only in it for the money anyway, are not to be trusted at all.
4. My friend Ellen is a huge snob about music. She says that rock-n-roll, a degenerate genre, is no better than elevator music. / According to her, rock-n-roll, a degenerate genre, is no better than elevator music.
5. Poor Joan seems to have grown crazier than ever. She now claims that her apartment was bugged by the Feds, who are listening to her every word. / Her apartment was bugged by the Feds, who are listening to her every word.
6. My brother Sid hates school. He says that he puts off his homework, a complete waste of time, to the last minute. / He puts off his homework, a complete waste of time, to the last minute.
7. I talked to an outlandish theater critic at a party. He told me that modern theater, which has been on the decline for years, is near its end. / According to him, modern theater, which has been on the decline for years, is near its end.
8. My kid sister Loni is obsessed with comic books. She says that a good graphic novel, man's greatest achievement, can keep her up reading until dawn. / A good graphic novel, man's greatest achievement, can keep her up reading until dawn.

### B Epithet materials

The experiment consisted of two factors. The slash (/) indicates which adjective or verb was used to create a Positive/Negative context, while the Intensifier is given in parentheses.

1. My classmate Sheila said that her history professor gave her a (really) high/low grade. The jerk always favors long papers.
2. My neighbor Maria said that her roommate has the (absolute) best/worst sense of humor. The idiot never stops talking.
3. My roommate Glen said that his uncle tells the (absolute) funniest/lamest jokes. The stooge can never get through a single one of them without giggling.
4. My sister Trudy said that her blind date showed up wearing an (incredibly) expensive/tasteless suit. The idiot spent a lot of money to impress her.
5. My buddy Steve said that his landlord plays the trumpet (very) well/badly. The twerp never takes a night off.
6. My co-worker Miranda said that our boss gave her a (super) generous/stingy Christmas bonus. The skinflint has always treated the pretty ones better.
7. My brother Ken said that his math tutor has been in (such) a great/terrible mood lately. The jerk is always nicer when he's paid in advance.
8. My friend Mike said that his housemate threw a (totally) fantastic/horrible party last weekend. The cretin always invites a lot of people.
9. My aide Sandy said that the good-looking bike messenger is always (so) sweet/rude to her. The creep tried to seduce her in the past.
10. My colleague Sarah said that the president of the company (totally) supported/denied her promotion. The jerk doesn't know the first thing about talent.
11. My sister Katie said that her art teacher (really) praised/criticized her painting in front of the whole class. The moron thinks anything in watercolor is a masterpiece.
12. My cousin Claire said that her landlord (totally) lowered/raised the rent last all of a sudden. The bastard is trying to take advantage of her.
13. My friend Carl said that his lab partner got the (absolute) best/worst score on the midterm. The twerp barely cracked a book all term.
14. My classmate Heidi said that the judge awarded her (such) a great/lousy prize at the science fair. The cretin was probably taking bribes.
15. My buddy Connor said that his boss bought (such) expensive/cheap computers for the office. The imbecile wants to impress the CEO at the next board meeting.
16. My sister Lyra said that her ex-boyfriend tried to take her to a (super) pricey/trashy restaurant on their anniversary. The cheapskate keeps trying to win her back.

### C Corpus examples

The following is example 2 from the appositives corpus (Potts and Harris 2009). An example of a non-speaker-oriented case is (23). The full corpus is downloadable from the Internet. The distribution includes the full corpus as well as an associated search tool and other supporting materials.

The assessors saw only the basic information and the authors' evidence. The other author annotations are included in the corpus distribution to facilitate other kinds of searching and categorization.

- *Basic information*
  - *Embedding predicate*: report
  - *Appositive*: ESPN reported on Sunday, April 11, that the Lightning, who have been playing in 10,400-seat Expo Hall, are exploring opportunities to move to either Atlanta or Minneapolis.
  - *Full document*: As their first season comes to a close, there are rumors swirling that the Tampa Bay Lightning just might become the Atlanta Lightning. Or the Minnesota Lightning. But they are just rumors, according to Lightning general manager Phil Esposito. ESPN reported on Sunday, April 11, that the Lightning, who have been playing in 10,400-seat Expo Hall, are exploring opportunities to move to either Atlanta or Minneapolis. But Esposito said there was no truth to the report. “We were disappointed with ESPN’s irresponsible comment,” Esposito said. “There is just no substance to the rumor.” Rumors have been swirling for the past 2 years regarding a move by Tampa Bay, which is presently discussing plans to build a new arena.
  - *Source*: 20\_newsgroups/rec.sport.hockey/53468
- *Authors’ annotations*
  - *Appositive scope*: Text
  - *Evidence*: The next sentence is “But Esposito said there was no truth to the report”. This clearly does not target the appositive content concerning the current stadium. Thus, the appositive is not embedded under *report*
  - *Evidence status*: Good
  - *Evidence category*: Separate assertion(s); Contradiction
- *Assessor 1*
  - *Evidence type*: Text
  - *Evidence status*: Good
  - *Notes*: Relies on world knowledge of what is plausibly being denied by Esposito; it is presumably open and easily verifiable information where the team is currently playing.
- *Assessor 2*
  - *Evidence type*: Text
  - *Evidence status*: Good

## References

- Aloni, M. (2000). *Quantification under conceptual covers*. Ph.D. thesis, Universiteit van Amsterdam, Amsterdam. Published in the ILLC Dissertation Series, 2001-1.
- Amaral, P., Roberts, C., & Allyn Smith, E. (2007). Review of *the logic of conventional implicatures* by Chris Potts. *Linguistics and Philosophy*, 30(6), 707–749.
- Anand, P. (2007). Re-expressing judgment. *Theoretical Linguistics*, 33(2), 199–208.
- Anand, P., & Nevins, A. (2004). Shifty operators in changing contexts. In K. Watanabe & R. B. Young (Eds.), *Proceedings of SALT 14* (pp. 20–37). Ithaca, NY: CLC Publications.
- Aoun, J., Choueiri, L., & Hornstein, N. (2001). Resumption, movement, and derivational economy. *Linguistic Inquiry*, 32(3), 371–403.
- Asher, N. (2000). Truth conditional discourse semantics for parentheticals. *Journal of Semantics*, 17(1), 31–50.

- Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics*. Cambridge: Cambridge University Press.
- Bach, K. (1999). The myth of conventional implicature. *Linguistics and Philosophy*, 22(4), 367–421.
- Bach, K. (2006). Review of Christopher Potts, ‘The logic of conventional implicatures’. *Journal of Linguistics*, 42(2), 490–495.
- Beaver, D., & Clark, B. Z. (2008). *Sense and sensitivity: How focus determines meaning*. Oxford: Wiley-Blackwell.
- Bergler, S. (1992). *The evidential analysis of reported speech*. Ph.D. thesis, Brandeis University.
- Boër, S. E., & Lycan, W. G. (1976). *The myth of semantic presupposition*. Bloomington, IN: Indiana University Linguistics Club.
- Büring, D. (2005). *Binding theory*. Cambridge: Cambridge University Press.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46.
- Constant, N., Davis, C., Potts, C., & Schwarz, F. (2009). The pragmatics of expressive content: Evidence from large corpora. *Sprache und Datenverarbeitung*, 33(1–2), 5–21.
- Corazza. (2005). On epithets qua attributive anaphors. *Journal of Linguistics*, 41(1), 1–32.
- Crawley, M. J. (2007). *The R book*. West Sussex, UK: Wiley.
- Cruse, D. A. (1986). *Lexical semantics*. Cambridge: Cambridge University Press.
- Culicover, P. W. (1992). English tag questions and Universal Grammar. *Lingua*, 88(1), 21–54.
- de Marneffe, M.-C., Rafferty, A. N., & Manning, C. D. (2008). Finding contradictions in text. In *Proceedings of the Association for Computational Linguistics 2008* (pp. 1039–1047). Columbus, OH: Association for Computational Linguistics.
- Emonds, J. E. (1976). *A transformational approach to English syntax*. New York: Academic Press.
- Farkas, D. F. (1992). On the semantics of subjunctive complements. In P. Hirschbühler (Ed.), *Romance languages and linguistic theory* (pp. 69–104). Amsterdam: John Benjamins.
- Giannakidou, A. (1999). Affective dependencies. *Linguistics and Philosophy*, 22(4), 367–421.
- Giannakidou, A., & Stavrou, M. (2008). On metalinguistic comparatives and negation in Greek. In *Greek syntax and semantics*. MIT Working Papers in Linguistics. Cambridge, MA: MIT Linguistics.
- Godfrey, J. J., & Holliman, E. (1993). *Switchboard-1 transcripts*. Linguistic Data Consortium, Philadelphia.
- Harris, J. A. (2009). Epithets and perspective shift: Experimental evidence. In M. Biezma & J. A. Harris (Eds.), *University of Massachusetts occasional papers in linguistics: Papers in pragmatics* (Vol. 39). Amherst, MA: GLSA Publishing.
- Harris, J. A., & Potts, C. (2009). Predicting perspectival orientation for appositives. In *Proceedings from the 45th annual meeting of the Chicago Linguistic Society* (Vol. 45). Chicago, IL: Chicago Linguistic Society.
- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59(4), 447–456.
- Kaplan, D. (1999). What is meaning? Explorations in the theory of *Meaning as Use*. Brief version—draft 1. Ms., UCLA.
- Karttunen, L. (1973). Presuppositions and compound sentences. *Linguistic Inquiry*, 4(2), 169–193.
- Karttunen, L., & Zaenen, A. (2005). Veridicity. In G. Katz, J. Pustejovsky, & F. Schilder (Eds.), *Annotating, extracting and reasoning about time and events*, No. 05151 in Dagstuhl Seminar Proceedings. Dagstuhl, Germany: Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany. <http://drops.dagstuhl.de/opus/volltexte/2005/314>.
- Kuno, S. (1987). *Functional syntax: Anaphora, discourse, and empathy*. Chicago: University of Chicago Press.
- Lakoff, G. (1966). Deep and surface grammar. Ms., Harvard University.
- Lasnik, P. (2005). Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy*, 28(6), 643–686.
- Lewis, D. (1969). *Convention*. Cambridge, MA: Harvard University Press. Reprinted 2002 by Blackwell.
- Lewis, D. (1975). Languages and language. In K. Gunderson, (Ed.), *Minnesota studies in the philosophy of science* (Vol. VII, pp. 3–35). Minneapolis: University of Minnesota Press.
- Manning, C. D. (2006). Local textual inference: It’s hard to circumscribe, but you know it when you see it—and NLP needs it. Ms., Stanford University.
- Marcus, M. P., Santorini, B., Marcinkiewicz, M. A., & Taylor, A. (1999). *Treebank-3*. Linguistic Data Consortium, Philadelphia.

- McCawley, J. D. (1982). Parentheticals and discontinuous structure. *Linguistic Inquiry*, 13(1), 91–106.
- McCawley, J. D. (1987). Some additional evidence for discontinuity. In G. J. Huck & A. E. Ojeda (Eds.), *Syntax and semantics*, (vol. 20: Discontinuous Constituency (pp. 185–200). Orlando, FL: Academic Press.
- McCawley, J. D. (1989). Individuation in and of syntactic structures. In M. R. Baltin & A. S. Kroch, (Eds.), *Alternative conceptions of phrase structure* (pp. 117–138). Chicago: University of Chicago Press.
- Miltsakaki, E., Prasad, R., Joshi, A., & Webber, B. (2004). The Penn discourse TreeBank. In *Proceedings of the language resources and evaluation conference*. Lisbon, Portugal.
- Oshima, D. (2006). *Perspectives in reported discourse*. Ph.D. thesis, Stanford University.
- Pollard, C., & Sag, I. A. (1992). Anaphors in English and the scope of the binding theory. *Linguistic Inquiry*, 23(2), 261–303.
- Potts, C. (2005). *The logic of conventional implicatures*. Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press.
- Potts, C. (2007). The expressive dimension. *Theoretical Linguistics*, 33(2), 165–197.
- Potts, C. (To appear). Conventional implicature and expressive content. In C. Maienborn, K. von Heusinger, & P. Portner, (Eds.), *Semantics: An international handbook of natural language meaning*. Berlin: Mouton de Gruyter.
- Potts, C., & Harris, J. A. (2009). Embedded appositives (annotated data). <http://christopherpotts.net/ling/data/embeddedappositives/>.
- Potts, C., & Schwarz, F. (2008). Exclamatives and heightened emotion: Extracting pragmatic generalizations from large corpora. Ms., UMass Amherst.
- Quang, P. D. (1971). English sentences without overt grammatical subject. In A. M. Zwicky, P. H. Salus, R. I. Binnick, & A. L. Vanek (Eds.), *Studies out in left field: Defamatory essays presented to James D. McCawley on the occasion of his 33rd or 34th birthday*, (pp. 3–10). Amsterdam: John Benjamins.
- R Development Core Team. (2008). *R: A language and environment for statistical computing*. organization R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, <http://www.R-project.org/>.
- Ross, J. R. (1973). Slifting. In M. Gross, M. Halle, & M.-P. Schützenberger (Eds.), *The formal analysis of natural languages* (pp. 133–169). The Hague: Mouton de Gruyter.
- Sauerland, U. (2007). Beyond unpluggability. *Theoretical Linguistics*, 33(2), 231–236.
- Saurí, R., & Pustejovsky, J. (2009). FactBank: A corpus annotated with event factuality. *Language Resources and Evaluation*, 43(3), 227–268.
- Schlenker, P. (2003). A plea for monsters. *Linguistics and Philosophy*, 26(1), 29–120.
- Schlenker, P. (2007). Expressive presuppositions. *Theoretical Linguistics*, 33(2), 237–245.
- Simons, M. (2007). Observations on embedding verbs, evidentiality, and presupposition. *Lingua*, 117, 1034–1056.
- Smith, C. S. (2009). Accounting for subjectivity (point of view). In R. P. Meier, H. Aristar-Dry, & E. Destruel (Eds.), *Text, time, and context: Selected papers of Carlota S. Smith, studies in linguistics and philosophy* (Vol. 89, pp. 371–393). The Netherlands: Springer.
- Taylor, K. A. (2007). Misplaced modification and the illusion of opacity. In M. O'Rourke & C. Washington, (Eds.), *Situating semantics: Essays on the philosophy of John Perry*, (pp. 215–250). Cambridge, MA: MIT Press.
- von Stechow, A. (2003). Binding by verbs: Tense, person and mood under attitudes. In M. Kadowaki & S. Kawahara (Eds.), *Proceedings of the 33rd north east linguistics society meeting*. Amherst, MA: GLSA.
- Wagner, M. (2004). Asymmetries in the syntax and prosody of verb-initial interpolated clauses in German. In S. Blaho, L. Vincent, & E. Schoorlemmer (Eds.), *Proceedings of CONSOLE XII*, (pp. 201–215). Leiden: The Student Organization of Linguistics in Europe.
- Wald, A. (1943). Test of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54(3), 426–482.
- Wang, L., Reese, B., & McCready, E. (2005). The projection problem of nominal appositives. *Snippets*, 10, 13–14.
- Zaenen, A., Karttunen, L., & Crouch, R. (2005). Local textual inference: Can it be defined or circumscribed? In *ACL workshop on empirical modelling of semantic equivalence and entailment*. Ann Arbor, MI: Association for Computational Linguistics.