

Processing presuppositions: dynamic semantics vs pragmatic enrichment.*

Emmanuel Chemla (Institut Jean Nicod & LSCP, France)

Lewis A. Bott (Cardiff University, UK)

To appear in *Language and Cognitive Processes*

Last modified: July 2011

Abstract: *One defining and yet puzzling feature of linguistic presuppositions is the way they interact with linguistic operators. For instance, when a presupposition trigger (e.g., realise) occurs under negation (e.g., Zoologists do not realise that elephants are mammals), the sentence is most commonly interpreted with the same global presupposition (elephants are mammals) as if negation was not present. Alternatively, the presupposition may be locally accommodated, i.e., the presupposition may become part of what is negated. In this paper, we develop and test two processing accounts of presupposition projection, the global-first model and the local-first model, inspired by dynamic semantics and pragmatic theories respectively. We tested these predictions using a verification task similar to Bott and Noveck's (2004) test of default models of scalar implicature. Across two experiments, using different materials and instructions, participants were faster to derive the global interpretation than the local interpretation, in contrast to the local-first model. We discuss the results in terms of dynamic semantics vs. pragmatic models of presupposition projection (e.g., Heim, 1983b vs. Schlenker, 2008).*

Keywords: *presupposition; processing; scalar implicatures; pragmatics; dynamic semantics*

Linguists and philosophers have long recognised that natural languages offer the means to distinguish between the main point of an utterance, its *assertive content*, and information that should be considered as background information for the participants to the conversation, so-called *presuppositions*. The large set of linguistic expressions or constructions that are said to “trigger” a presupposition includes definite descriptions (*the*), change of state predicates (*stop, start, continue*), additive particles (*too*), pseudo-clefts (*it is X who...*), or, importantly for this paper, factive verbs such as *realise, discover, know, regret, ignore*. For example, each of the sentences below has an *assertive* component and a *presupposition* component.

- (1)
 - a. The king of Moldavia is wise.
 - b. John stopped smoking.
 - c. Mary got married too.
 - d. It is Helen who killed Bob.
 - e. Michael discovered Jenna was having an affair.

(1a) *asserts* that the king of Moldavia is wise, but *presupposes* that Moldavia is a monarchy; (1b) *asserts* that John does not smoke, but *presupposes* that he used to; (1c) *asserts* that Mary got married, but *presupposes* that someone else also got married; (1d) *presupposes* that Bob was killed, and *asserts* that the killer was Helen; (1e) *asserts* that Michael became aware that Jenna was having an affair, but *presupposes* that she was. In this paper we derive and test processing predictions about how presuppositions interact with linguistic operators such as negation.

* The authors are grateful to Jack Tomlinson, Bart Geurts and to the editor and reviewers for *Language and Cognitive Processes* for their input. This work was supported by a ‘Euryi’ grant from the European Science Foundation (“Presupposition: A Formal Pragmatic Approach”) and by the Cardiff University International Collaboration Fund.

In the next section we describe the general phenomenon under investigation, known as the *projection problem*, followed by linguistic accounts of the problem. We then show that natural cognitive implementations of these theories lead to different processing predictions. Finally we present two experiments which offer the first behavioural results that address these issues and show how to adjudicate between these processing accounts.

Presupposition projection

Presuppositions behave differently from assertions in several ways. One such difference is how they behave when presupposition triggers are embedded in more complex sentences, such as under negation, or modals. This complex interaction between presupposition and linguistic operators has been studied as the *projection problem for presupposition*. Let us start with an example to illustrate. Given what may be assumed about negation, presuppositions interact with negation in a surprising way. Consider examples (2) and (3). Because (3) is the negation of (2), the assertive components of the two sentences are opposite. But they have the same presupposition, that elephants are mammals. The same would be true for each of the examples in (1): adding a negation would not affect their presupposition. More generally, linguistic operators – negation, conjunction (e.g., *and*, *but*, *because*), disjunction (e.g., *or*), quantifiers (e.g., *some*, *many*, *all*), etc. – do not act on presuppositions as it would be expected from their action on the assertive component of a given phrase.

- (2) Zoologists realise that elephants are mammals.
 - Presupposition: *elephants are mammals*.
 - Assertion: *Zoologists are aware that elephants are mammals*.
- (3) Zoologists do not realise that elephants are mammals.
 - Presupposition: *elephants are mammals*.
 - Assertion: *Zoologists are NOT aware that elephants are mammals*.

This interaction is studied as the *projection problem*, which is, as Heim (1983a) puts it, “the problem of predicting the presuppositions of complex sentences in a compositional fashion from the presuppositions of their parts”. As illustrated above, presuppositions in complex sentences (e.g., with a negation) do not behave in a straightforward manner and the projection problem seeks to explain this behaviour. At issue is the compositionality of language, i.e. the standard way in which we construct sentence meaning from the meaning of individual parts of the sentence. Not surprisingly, such a fundamental problem has attracted a lot of attention since it was discovered but, more interestingly, there is still no accepted account of the phenomena (see Beaver and Geurts, to appear, for a recent overview). In the next section we discuss two types of influential linguistic accounts of presupposition projection and try to give a sense of the current theoretical considerations that animate the debates.

Linguistic accounts of presupposition projection

Dynamic semantics models (e.g., Heim, 1983b) claim that the meaning of any expression has two components: one for presupposition and one for assertion. A linguistic operator may thus act differently on the presupposition and on the assertion of the linguistic material *X* to which it applies. The projection problem is solved by setting appropriately the presuppositional component of the relevant operators. For example, negation is given a lexical entry as in (4), which recovers the pattern we described above: negation is active only at the level of the assertion.

- (4) negation [*X*] = Presupposition of *X*

and
not [Assertion of *X*]

In this view, presupposition is a *semantic* phenomenon in the sense that presuppositions of sub-sentential linguistic constituents are computed along with other grammatical, recursive computations of meaning.

In recent years, this semantic approach has been criticised¹ and alternative *pragmatic* models of presupposition projection have been developed to offer new formal solutions (of particular importance for our purposes are Simons, 2004, Schlenker, 2008, Chemla, 2008a, 2010, Abusch, 2010, Romoli, 2011; see Schlenker, 2010 for an overview). In these views, negation is given a more traditional, non-dynamic lexical entry, i.e., a lexical entry which does not distinguish between presuppositional and non-presuppositional aspects of the meaning of the constituent it applies to:

(5) negation [*X*] = not [Meaning of *X*]

Presuppositions are thus silent at the semantic level at which, e.g., negation applies; they come into play when general rules of conversation are considered, and these rules become relevant only after the recursive computation of meaning has been terminated. The differing behaviour of assertions and presuppositions under operators is therefore not determined at a lexical level (as in (4)), but by conversational pragmatics.

There are many different versions of the semantic and pragmatic theories that we have outlined above and our experiments do not relate directly to any one of them in particular. Instead, our experiments seek to distinguish between two classes of theories: those theories that make presupposition projection obligatory and necessary (mostly *semantic* theories), and those theories that make it a process occurring after other more fundamental semantic computations (mostly *pragmatic* theories). We present below an overview of Schlenker (2008), to give a sense of the pragmatic considerations and tools that may be recruited in the second class of theories.

A recent example of a pragmatic theory

Schlenker (2008) seeks to explain presupposition projection as manner implicatures, in Grice's sense (Grice, 1967). The core idea is that the information conveyed by presuppositional phrases is dense, e.g., *realise* conveys both that the proposition expressed in its complement is true and that its subject argument holds it as true. Hence, phrases or sentences containing a presupposition trigger, e.g., (6)a, are claimed to evoke more 'articulated', albeit semantically equivalent alternative phrases, e.g., (6)b.² Now, if the

¹ While being highly influential, dynamic semantics has been criticised on conceptual grounds because the result it aims to achieve is stipulated in enriched lexical entries, thus failing to provide an *explanatory* solution to the projection problem and merely offering a framework well-suited for a systematic *description* of the facts (see e.g., Soames, 1989, for early criticism). In fact, not only would negation have to be revisited and receive a new lexical entry as above, but all other standard meanings of otherwise standard operators like conjunction, disjunction, quantifiers would have to be modified on a case by case basis to fit the presupposition data.

² For readers familiar with quantity or scalar implicatures, (6)b plays the role of what would be the stronger alternative. The claim here is that presuppositions are *manner* implicatures: the alternative is preferable although not because it is more informative.

speaker chose not to utter the more articulated form, e.g., 6(b), it can be assumed that the current conversational context is such that there is no advantage in using the more articulated form. In our example, this inference amounts to the fact that the truth of the information in the complement of the verb *realise* is agreed on. In short, an utterance of the condensed version (6)a will trigger the implicature that the additional explicitness of its competitor (6)b is unnecessary, which happens if participants to the conversation agreed on the truth of the alleged presupposition.

- (6) a. Zoologists realise that elephants are mammals.
 b. Elephants are mammals and Zoologists realise that <elephants are mammals>.

A detailed formalisation of this competition between explicitness and briefness leads to a general pragmatic explanation of presupposition projection.³ In general, the condensed version of the sentence will only be acceptable when the more explicit form would not make a different contribution than the shorter one. When applied to sentences involving negation, this principle dictates that presuppositions escape negation, without lexical specification about negation.

We have presented Schlenker's theory as an example of how pragmatic accounts explain presupposition projection. However, the details of the theory are not important for the present purposes. What matters is that under pragmatic views, presuppositions come out as a negotiation between different maxims, exactly as for other types of implicatures (e.g., scalar implicatures, e.g., Horn, 1972, are explained via maxims of quantity and relevance). This means that the explanations arise without any specific stipulations about the lexical properties of negation or any other linguistic operator (see the criticisms of dynamic semantics presented in Footnote 1), and that presuppositions only arise at a stage at which pragmatic considerations become relevant, that is, after the literal meaning of the sentence has been generated.

We next describe how the semantic and pragmatic theories derive multiple interpretations of presupposition sentences under negation, a phenomenon we use to test between processing instantiations of the different types of theories.

Presupposition under negation, another reading: local accommodation or literal meaning?

Both dynamic semantics and pragmatic accounts predict that a negative sentence like (3) has the same presupposition as its positive counterpart (2). However, under certain circumstances, the presupposition of a negative sentence seems to disappear. This is apparent in discourses like (7):

- (7) Of course, zoologists do not realise that elephants are birds, because elephants are *not* birds!

³ The interested reader is referred to the original theory but we provide some more details here. The theory predicts that a sentence such as $F(pp')$, where a phrase pp' with presupposition p and assertion p' is embedded in an environment $F(\dots)$, will presuppose the following: $\forall x, F(p \text{ and } x) \Leftrightarrow F(x)$. The idea is that the condensed sentence $F(pp')$ is acceptable only if a more explicit, or articulated sentence of the form $F(p \text{ and } x)$ (the left-hand side of the equivalence) would not make a different contribution than a shorter one such as $F(x)$ (the right-hand side of the equivalence). (The quantification over x s can be understood as an abstraction away from the assertion or from the material following the occurrence of the presupposition). For instance, the presupposition of *negation* (pp') would be $\forall x, \text{not } (p \text{ and } x) \Leftrightarrow \text{not } (x)$. In particular, with x instantiated as the tautology T , this leads to $\text{not } (p \text{ and } T) \Leftrightarrow \text{not } (T)$, i.e. p is true. Hence, the result that a sentence and its negation have the same presupposition can be retrieved.

What appears to be happening in (7) is that both the assertion and the *elephants are birds* presupposition are being negated. The presuppositional clause is therefore interpreted as meaning something like *it is not the case that (elephants are birds and zoologists know this)*. This contrasts with the standard behaviour in which the presupposition escapes negation, as in (4). Negative sentences as in (3) or (7) can therefore be said to have a non-presuppositional interpretation in which, descriptively, the presupposition remains trapped under negation, as schematised in (8):

$$(8) \quad \text{negation [Z realise p]} = \text{not [p and Z believe p]}$$

In dynamic semantics, this non-presuppositional interpretation is explained by means of a process called *local accommodation*. According to this process, the presupposition of a constituent *X* may be treated as a genuine part of the assertion if something goes wrong with the application of the regular dynamic negation (4) to *X*. In (7), the application of the routine rule (4) leads to a contradiction in which the sentence would convey that *elephants are birds*, despite the fact that everyone knows that it is not true to begin with and that the opposite is actually asserted right after. To resolve this situation, the presupposition of *X* is cancelled, i.e., it becomes part of the normal asserted meaning. Descriptively, then, the presupposition is interpreted locally in (7) (i.e., under negation where it is triggered) and we will therefore refer to this interpretation as the *local reading*. The other interpretation, in which presupposition escapes negation and is interpreted at the level of the whole sentence, will be referred to as the *global reading*. This terminology is summarised in Table 1:

Table 1. Different readings for negative sentences with a presupposition, and schematic predictions of different approaches

	Global reading (Presupposition escapes negation and is interpreted at the <i>global</i> level)	Local reading (Presupposition is interpreted under negation, at the <i>local</i> level where it was triggered)
Negation of S	PRESUPPOSITION OF S AND NOT [ASSERTION OF S]	NOT [PRESUPPOSITION OF S AND ASSERTION OF S]
Zoologists do not realise that elephants are mammals	Elephants are mammals AND NOT[Zoologists believe that]	NOT [Elephants are mammals AND Zoologists believe that]
Semantic accounts	Core reading → Altered reading	
Pragmatic accounts	Pragmatic reading ← Literal meaning	

In the pragmatic approach, the non-presuppositional, or local, reading described in (8) is straightforwardly explained: it corresponds to the literal meaning of the sentence as predicted by (5), the standard lexical entry for negation, before pragmatic processes apply.

Hence, it is predicted to be an available reading for the sentence, although one that should be superseded when pragmatic processes come into play and enrich it into the global reading.

Processing models

Both dynamic semantics and the pragmatic approach can accommodate the different readings available for negative sentences but, crucially, they suggest different processing models. The dynamic semantics account suggests that the initial meaning of presupposition sentences with negation is the global interpretation, i.e., the presupposition escapes negation. This follows from the dynamic lexical entry for the negation operator (as in (4)). Only if this interpretation is judged unacceptable is the local interpretation (as in (8)) derived.⁴ Consequently we refer to the processing instantiation of the semantic account as the global-first model.

To illustrate the predictions of the global-first model, consider sentence (3), *Zoologists do not realise that elephants are birds*. Under the global-first model, the presupposition trigger, *realise*, causes the negation operator to apply differently to the assertion and the presupposition, as in (4). This results in the global interpretation, something like *elephants are birds and Zoologists realise that elephants are birds*. Because the presupposition *elephants are birds* is false in this case, the listener may however, look for a more charitable interpretation of the sentence. One solution is to weaken the role of the faulty presupposition and to treat it as an asserted part of the sentence. This would lead to the local interpretation, *it is not the case that [elephants are birds and Zoologists believe so]*.

The pragmatic account, on the other hand, assumes a standard lexical entry for negation, shown in (5), which means that the local interpretation is derived first: it is the literal meaning of the sentence in this view. The global interpretation thus comes out through pragmatic processes that take as input the local reading, and hence should unfold as a later (re)interpretation. The pragmatic processing account will therefore be referred to as the local-first model. Consider (3) again, *Zoologists do not realise that elephants are birds*. Under the local-first model, deriving the literal meaning of the sentence involves applying the negative operator to the sentence as a whole (see (5)). This results in the local interpretation, something like, *it is not the case that [elephants are birds and Zoologists believe so]*. However, pragmatic maxims may then be applied to the literal meaning of the sentence (i.e., the local interpretation), which results in the global interpretation.

Our main aim in this article is to derive and test processing models of presupposition projection. In doing so we make two standard assumptions about processing and pragmatics. These are (i) that the literal meaning of the sentence is the input to pragmatic procedures, and (ii) that the literal meaning is accessible prior to the application of pragmatic procedures. Given these assumptions, we claim that the pragmatic and dynamic semantics representational theories naturally lead to two distinct alternatives, the global-first and local-first hypotheses. It may be possible to defend different processing implementations of the dynamic and pragmatic theories, but we believe that the burden of the proof would then be on the independent arguments for such re-interpretations. We return to this issue in the General Discussion.

From a methodological perspective, we capitalise on the fact that a similar situation has been identified and studied in the domain of scalar implicatures (e.g., Horn, 1972), another

⁴ This implementation is particularly warranted in so-called *cancellation* theories. Such theories solve the projection problems by means of late local accommodation processes, which wipe out problematic global presuppositions coming from a primary derivational step and help deliver contradiction-free presuppositional outputs, see e.g., Gazdar (1979).

phenomenon at the interface between semantics and pragmatics. Our experiments borrow from this literature and test the presupposition projection models using a paradigm with a logic similar to Bott and Noveck (2004), one of the pioneer studies in the domain of scalar implicatures which has now led to numerous refinements.

Experiment 1

Our approach was to compare processing times for global and local interpretations of sentences like (3). If people derive global interpretations and then potentially revert to local interpretations, as in the global-first account, processing times should be shorter for global interpretations than local interpretations. Conversely, the local-first account suggests that there will be an early derivation of the local, bare semantic interpretation of the sentence from which the global interpretation is constructed: local interpretation times should be shorter than global interpretation times. (Note that the amount of time needed to derive either of the interpretations individually is unimportant, as is the process by which either interpretation is derived; all that matters is the serial nature of the local- and global-first hypotheses.)

Participants completed a sentence verification task. The experimental sentences all involved a negated factive verb (*realise*) with a complement that generated a false presupposition (e.g., *elephants are birds*). The experimental sentences therefore had the following global and local interpretations (see also Table 1):

- (9) Zoologists do not realise that elephants are birds.
- a. Global: [*Elephants are birds*] and not [*zoologists believe so*] (false)
 - b. Local: NOT [(*Elephants are birds*) AND (*zoologists believe so*)] (true)

Global accommodation interpretations should therefore generate a false response whereas local accommodation interpretations should generate a true response. All things being equal, a local-first account consequently predicts shorter response times to true responses than false responses whereas a global-first account predicts the reverse pattern. We included experimental sentences, as described above, and four types of control sentences. These are shown in Table 2. Control sentences (b) and (c) were needed to ensure that participants were unable to predict the correct response prior to reading the final word (participants were given a cover story designed in particular to make the (c) condition unambiguously true) and control sentences (d) and (e) were needed to estimate any bias against false responding in general (see e.g., Clark & Chase, 1972).

Method

Participants

Thirty-three Cardiff University students participated for course credit. Three participants were removed for poor performance on the control sentences (see below).

Design

Each sentence was formed using *Geographers* or *Zoologists* as the subject, a factive verb with negation (*do not realise*) or a non-factive verb (*tell*), and a proposition about categories as the complement of the verb (e.g., *elephants are mammals* or *elephants are birds*), as in (10) below.

- (10) {Zoologists / Geographers} {do not realise / were told} that *subcategory* are *supercategory*.

We generated 60 place names and 60 animals as the subordinate category member of the category proposition. These exemplars formed the basis of each item in the design. The experimental sentences were generated using an exemplar and an incorrect supercategory, combined with the appropriate professional (zoologists or geographers), as in (9). Four control versions of each item were formed using the same exemplar but with a different subject or a different superordinate category (i.e., the correct superordinate) to obtain unambiguously true and false sentences, with the presuppositional phrase (*do not realise*) and without a presuppositional phrase (*were told*). Table 2 shows the five versions of the *elephant* item.

Table 2. Example sentences and observed accuracy.

Condition	Example sentences (Exp 1)	Expected answers (Exp 1)	Observed “true” proportions (Exp 1)	Observed “correct” proportions (Exp 2)
(a)	Zoologists do not realise that elephants are reptiles.	True or False	.38 (.32)	.36 (.31)
(b)	Zoologists do not realise that elephants are mammals.	False	.12 (.073)	.14 (.09)
(c)	Geographers do not realise that elephants are mammals.	True	.85 (.094)	.83 (.11)
(d)	Zoologists were told that elephants are mammals.	True	.93 (.052)	.91 (.07)
(e)	Zoologists were told that elephants are reptiles.	False	.11 (.079)	.11 (.10)

Note. Condition (a) provides the experimental sentences. Conditions (b), (c), (d) and (e) are control sentences. Standard deviations in parentheses.

No subordinate category member was used more than once but superordinate category members were used multiple times. There were four superordinate geographical categories (Africa, Asia, Europe, America) and six superordinate zoological categories (birds, dogs, fish, insects, mammals and reptiles). Each superordinate appeared equally often across conditions.

Items were assigned to five counterbalancing lists (distributed equally among participants) so that all items appeared equally often in each condition, but no participant saw the same exemplar twice. In all, each participant saw 24 items in each condition.

Procedure

Participants were given a cover story to remove ambiguity about the situation, controlling especially the knowledge of zoologists and geographers. The cover story described an alien invasion of Earth in which different groups of aliens were trained to have specialist knowledge of Earth geography but no knowledge of Earth zoology (the geographers) or *vice versa* (the zoologists). This scenario allowed us to construct unambiguously true and false control sentences, with and without presupposition triggers. Participants also went through a

training phase in which they judged 24 control sentences and received feedback on their responses. No experimental sentences were presented during the training phase and participants did not receive feedback during the main part of the experiment.

Sentences were presented one word at a time in the centre of the screen. Each word was presented for 200ms except for the last word, which remained on the screen until the participant made their response.

Results

Data treatment

We removed three participants who scored less than .75 proportion correct on the control conditions. Responses with RTs greater than 10s were removed as outliers (18 out of 3600 data points) and the RT data were log-transformed to reduce positive skewness and inhomogeneity of variance. All reported means and standard deviations correspond to the raw data.

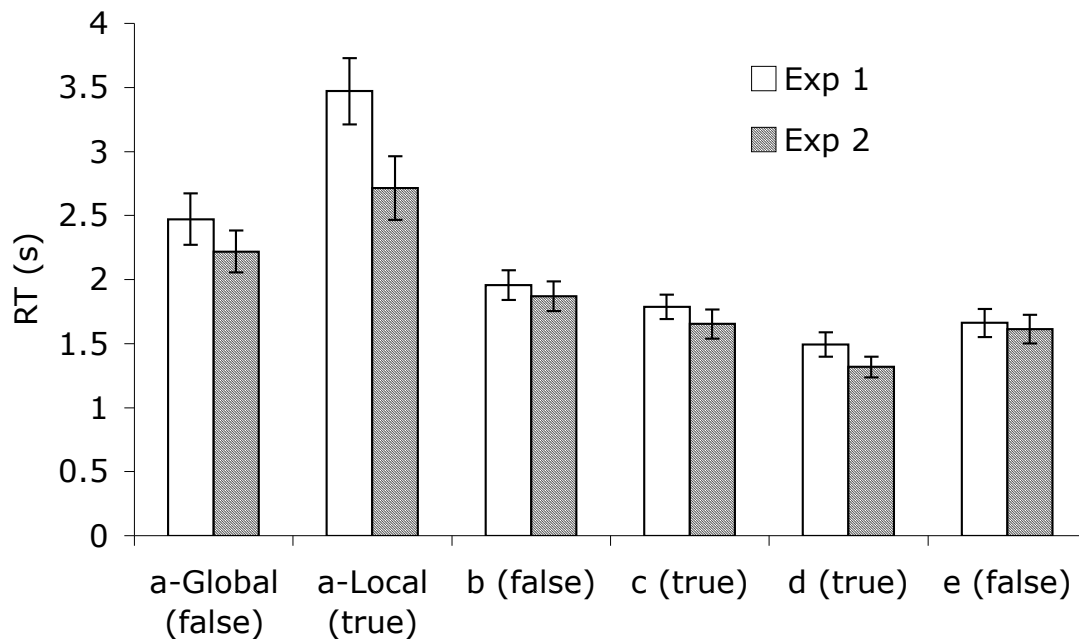
Choice proportions

Table 2 shows response proportions for all five conditions. Responses to control sentences were accurate $M = .89$, $SD = .047$, and demonstrate that participants had no difficulty understanding the cover story. Consistent with the co-existence of global and local derivations, the experimental sentences prompted a high degree of bivocality. Whilst there was a slight bias away from the local interpretation, $M = .38$, there was significantly greater variation in the experimental sentences than any of the control sentences, F^2 's > 40 , p 's $< .005$.

Response Times

Figure 1 shows the RTs as a function of the sentence type and the response type for the experimental sentences. Global responses (false) seem faster than local responses (true), consistent with a global-first account. We analysed this difference in two ways. First, we compared global and local RTs within each participant (and item). This analysis revealed that global interpretations to the experimental sentences were faster than local interpretations, $M = 2.47$ s ($SD = 1.37$) vs $M = 3.47$ s ($SD = 1.08$) $t_1(26) = 4.78$, $p < .005$, $\eta^2 = .47$ (three participants responded univocally and were therefore excluded from this analysis), $t_2(102) = 4.37$, $p < .005$, $\eta^2 = .16$ (17 items were excluded for the same reason). Second, we classified participants (and items) as local or global responders. Participants were ranked on the proportion of local responses they made to experimental sentences. The top half of participants were then classified as local responders and the bottom half as global responders. Consistent with the within-subject analysis, false responses from the global responders were faster than true responses from the local responders, $M = 2.75$ ($SD = .55$) vs $M = 3.23$ s ($SD = 1.11$) $t_1(28) = 3.86$, $p < .005$, $\eta^2 = .35$, $t_2(118) = 7.49$, $p < .005$, $\eta^2 = .32$. Overall then, the results are consistent with the global-first account.

Figure 1. Response time as a function of sentence type.



Note. Letters a, b, c, d, and e, correspond to the sentence types shown in Table 1. Condition (a) is broken down by response choice but all other conditions show RTs to correct responses only. Error bars refer to standard errors.

No comparable effect of answer (true vs false) was found on the RTs for the control conditions. Yet, one potential explanation for the slow local responses is that participants were slowed down by the *inconsistency* between the veracity of the embedded proposition (false) and the veracity of the local interpretation (true). We tested this by comparing local RTs to the experimental sentences against correct responses to (b) control sentences, in which the veracity of the embedded proposition was also contrary to the veracity of the correct response choice (false). Local interpretation RTs to (a) were significantly longer than correct RTs to (b), however, $M = 3.49\text{s}$ ($SD = 1.37$) vs $M = 1.95\text{s}$ ($SD = 0.63$), $t_1(27) = 7.24$, $p < .0005$, $\eta^2 = 0.66$, $t_2(102) = 9.15$, $p < .0005$, suggesting that long local RT cannot be due entirely to the inconsistency between embedded proposition and response choice.