Presupposition Satisfaction, Locality and Discourse Constituency

Christina S. Kim

Abstract The current study investigates presupposition-satisfying dependencies from the point of view of discourse processing. Using the presupposition trigger *also* as a case study, I ask to what extent the distance spanned by the trigger and the prior discourse content that satisfies the presupposition of *also* influences comprehenders' interpretation of the discourse — specifically, whether comprehenders have a bias toward satisfying presuppositions using material in the discourse that is closer rather than more distant.

Two offline experiments and one Visual World eye-tracking experiment provide evidence in favor of a locality bias in presupposition satisfaction. The findings support sensitivity to both linear distance and distance in terms of hierarchically structured discourse representations, consistent with an interpretation system that tracks both structure-insensitive information about discourse mention, and structured representations of larger discourse units.

1 Introduction

Sentences with *also* (1-a) are standardly assumed to entail their propositional content (1-b) and presuppose a distinct proposition which differs from the sentence's propositional content in the value of the focused constituent (Horn, 1969; Karttunen and Peters, 1979; Rooth, 1985; Atlas, 1991). That is, at least one alternative in *A* (1-c) is true, where the alternatives are restricted by the context.[1]

(1) a. Andy also bought some NECTARINES.
b. Andy bought some nectarines.

Christina S. Kim
University of Kent, Canterbury e-mail: c.s.kim@kent.ac.uk

[1] Assume the context restricts the set of possible alternatives to just the salient ones, where *salient* can be understood as *likely to be inferred by the addressee* (see e.g. Kim (2012)).

c. A = {Andy bought bread, Andy bought some celery, Andy bought a croissant, ...}

Thus, the discourse-final sentence (2-f) entails that Andy bought nectarines and presupposes that he bought something other than nectarines. If not already contextually entailed, presuppositions must be accommodated (Lewis, 1979) as background.

(2) a. The roommates often go to the farmer's market together.
b. Beth always buys bread.
c. Andy usually buys some celery.
d. His doctor told him he needs to eat more vegetables.
e. Today Andy treated himself to a croissant.
f. He also bought some NECTARINES.

From the point of view of processing, readers have been shown to experience processing difficulty at the presupposition trigger if the presupposed information has not been processed earlier in the discourse (Moulton, 2006; Schwarz, 2007). Such findings can be seen as characterizing the conditions under which presupposed content be understood in discourse as intended by the speaker. However, the question of which of multiple viable interpretations — different ways for a presupposition to be satisfied — is preferred by the listener has been relatively understudied.

The current study addresses the question of preference by viewing a presupposition trigger like *also* as forming a dependency with the discourse content that satisfies the presupposition, analogous to filler-gap or anaphoric dependencies (van der Sandt, 1992). By taking this view, we can ask whether such presupposition dependencies pattern like other kinds of dependencies formed during language comprehension, and are sensitive to factors known to constrain dependency formation at the sentential level. Specifically, locality (Hawkins, 1994; Gibson, 2000) has been shown to influence sub-sentential processing complexity. Here, I ask whether a similar bias in favor of local dependencies operates at the discourse level, using presupposition triggers and the discourse content they associate with as a case study. Asking this question opens up a way to address a related question about the mental representation of discourse — whether the relevant sense of locality is linear distance or distance relative to hierarchically-structured constituents (Webber and Joshi, 1998; Kehler, 2002).

The following sections outline some relevant prior findings that support structured representations of discourse (section 2), including evidence from language comprehension and production that discourses hypothesized to have different structures are also processed differently (section 2.1). Section 3 reviews evidence for a processing bias in favor of local dependencies. Two offline experiments assessing comprehenders' preferences for presupposition interpretation are presented in section 4. Section 5 presents a Visual World eye-tracking experiment (Tanenhaus et al., 1995) that tracks the online competition of interpretations representing linear and hierarchically-structured locality. Section 6 concludes.

2 Mental representations of the discourse

There is considerable evidence that discourse meaning cannot be adequately described using just the machinery of subsentential semantics and syntax. It is especially notable that research on discourse structure has been driven by different motivations and conducted using very different methodologies in neighboring fields of study. There is a rich history of research on narrative and text processing in psychology (Levin and Moore, 1977; Kintsch and van Dijk, 1978; Beaugrande and Colby, 1979; Schank et al., 1982; Morrow, 1985; Gee and Grosjean, 1984; Graesser and Singer, 1994; Simner and Pickering, 2005) as well as artificial intelligence and machine learning (Hayes, 1977; Cohen and Perrault, 1979; Litman and Allen, 1987; Hovy, 1993). A temporally parallel development moved formal linguists from a static view of sentential meaning toward context-update models of discourse, which take into account the sensitivity of linguistic meaning to a representation of the context that continually updates as a discourse progresses (Heim, 1982; Groenendijk and Stokhof, 1984; Kamp and Reyle, 1993).[2]

The major approaches to discourse structure differ along two dimensions that are relevant here. First, approaches differ in the types and number of relations that can hold between discourse units. In a number of prominent theories, the bulk of the explanatory work is done by a (presumably universal) limited inventory of discourse relations (Hobbs, 1979; Mann and Thompson, 1988; Kehler, 1995, 2002; Knott and Sanders, 1998; Wolf and Gibson, 2005).[3] While none of these theories in principle disallows structured representations of discourse, the emphasis remains on linear relations between adjacent segments (Figure 1).[4]

Existing approaches also differ in the extent to which discourse representations encode hierarchical and/or non-local structural relations among discourse units (McKeown, 1985; Polanyi and van den Berg, 1996; Webber and Joshi, 1998; Asher and Lascarides, 2003). These theories typically also feature an inventory of possible relations, but the characterizations of different relations include reference to their structural properties; for example, *Narration* is a coordinating relation in Segmented Discourse Representation Theory (Asher and Lascarides, 2003), while *Topic* is a subordinating one. Of primary importance here is that such structure-based theories attempt to explain what makes larger pieces of discourse cohere, by making reference to organizing principles like topichood.

A subclass of structure-based theories that emphasizes a particular organizing principle is the question-based theories of discourse, such as Roberts (1996)'s Question Under Discussion framework (Grosz and Sidner, 1986; van Kuppevelt, 1995;

[2] See also Hintikka (1976); Lewis (1979) for early work on dialogue/discourse-based approaches to meaning.

[3] See Sanders et al. (1992, 1993) for a taxonomy of coherence relations in terms of a small set of cognitive primitives, such as causality.

[4] This is primarily in reference to the coherence relations described in Hobbs (1979) and Kehler (2002). Rhetorical Structure Theory (Mann and Thompson, 1988) features both a large inventory of relations and the possibility of forming dependencies between non-atomic discourse constituents, producing hierarchical discourse structures.

Büring, 2003; Farkas and Bruce, 2009; Ginzburg, 2012). These approaches have in common the claim that discourses are organized around (often implicit) questions, or to use more general language, goals. A contribution to a discourse is appropriate, or coherent, if it provides a partial answer to the question of the current discourse; because partial answers include answers to sub-questions (and sub-sub-questions), discourse representations are inherently hierarchical (as illustrated in Figure 2). These approaches constitute a type of structure-based theory, but importantly adopt a view that questions (or goals) are what organize discourses.

2.1 Processing consequences of discourse structure

Importantly for current purposes, discourse factors have been shown to clearly influence interpretation in language processing. Explanations of coreference and binding (Grosz et al., 1995; Stevenson et al., 2000; Arnold and Griffin, 2007), syntactic ambiguity resolution (Frazier and Clifton Jr., 1996; Traxler et al., 1998; van Berkum et al., 1999; Hemforth et al., 2000), and ellipsis (Kehler, 2000; Ginzburg and Sag, 2000; Cooper and Ginzburg, 2002; Frazier and Clifton Jr., 2006; Kim and Runner, 2009) (among other phenomena) have appealed to properties of discourse relations or structures when interpretations vary in a way that syntactic or semantic properties cannot reliably predict.

A separate body of work in language processing collectively shows that discourse comprehension is influenced by aspects of discourse structure that go beyond sentence-level syntax and semantics. Studies in text processing (Sanders et al., 1992; Murphy and Shapiro, 1994) demonstrated that memory for discourses is enhanced by integrating a text into a larger discourse structure, and by manipulating the task associated with reading a text, suggesting that the comprehension of discourse content is affected by goal structure during memory encoding. In addition, the ability to

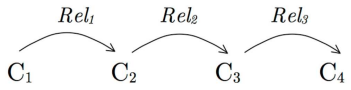

Fig. 1 Linear representation of discourse relations.

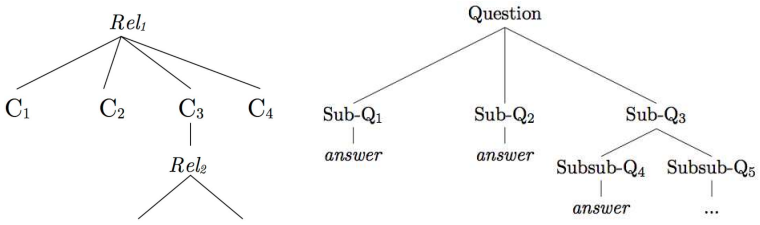

Fig. 2 Hierarchical representations of discourse.

detect contradictions, anomalies, and inconsistencies relative to discourse-level inferences is one of the signatures of depth of comprehension (Glenberg et al., 1982; Otero and Kintsch, 1992; Graesser and McMahen, 1993).

In online discourse processing, reading times show facilitation when a discourse resolves in a way that is expected based on prior discourse content or structure (for example, the discourse sets up an expectation for a causal continuation), and slow-downs when the actual continuation is an unexpected one. Bicknell and Rohde (2009) showed in a reading time study that comprehension in sentences containing ambiguously-attached relative clauses is influenced by the discourse relation that holds between the matrix and subordinate clauses.[5]

In language production, as well, language users appear to keep track of information related to discourse relations, and use that information to guide subsequent productions. For example, Simner and Pickering (2005) showed in a story continuation study that narratives containing information about the cause of an event generate more continuations related to consequences of the event, and narratives containing information about an event's consequences are followed by more continuations related to the event's cause. Rohde (2008) (Rohde et al., 2007; Kehler et al., 2008) show similar effects in sentence completion studies, where pronominal reference was influenced by the discourse relation that linked the clause containing the pronoun and the antecedent clause.

3 Locality effects in sentence processing

The research reviewed above demonstrate systematic processing reflexes of features of the discourse, such as discourse relations; this type of evidence tells us what kinds (and what grain) of information need or need not be encoded in the mental representation of the discourse. However, it does not tell us why or how those kinds of information should matter from the perspective of facilitating processing, or making it more efficient — that is, it does not tell us anything mechanistic. A body of research in sentence processing investigates dependency formation at the sentential level with such questions in mind.

The notion of dependency length plays a central role in proposals like Gibson (1998, 2000) and Hawkins (1994, 2004). Gibson's Dependency Locality Theory (DLT), in particular, links the lengths of dependencies (e.g. between a verb and its arguments) within a sentence to complexity, as reflected by processing time. A number of previously documented empirical phenomena in sentence processing receive a unified explanation in terms of the lengths of the dependencies they require comprehenders to construct. For example, King and Just (1991) (among others) showed that object-extracted relative clauses were more complex to process than their subject-extracted counterparts. The DLT explains this asymmetry in terms of two sources of complexity: a *storage cost*, which increases as the distance across which syn-

[5] For related studies on discourse effects in clause-level processing, see Millis and Just (1994); van Berkum et al. (1999); Roland et al. (2008); Rohde et al. (2011).

tactic predictions of previous words must be maintained in memory increases, and an *integration cost*, incurred when a dependency is formed with a preceding word, which increases as the distance spanned by this dependency increases. In the case of subject-extracted relative clauses, the verb in the relative clause is close (often adjacent) to the relativizer (e.g. *who*); the analogous dependency in a relative clause with an object gap is longer — minimally, it must be maintained across the subject argument — and as such incurs a higher integration cost. Related work by Temperley (2007) has since shown using corpora that production data also reflect a bias to minimize dependency lengths. Hawkins (1994, 2004) ties processing complexity to universals related to constituent ordering universals.

Part of the appeal of such theories is that their characterization in terms of notions like dependency and distance makes them in principle very broad-coverage. And since a theory about dependency length minimization crucially does not make reference to dependencies among verbs, subjects, and objects, we can ask whether other kinds of dependencies involved in language interpretation are *this* kind of dependency. The experiments presented here ask this question for dependencies triggered by presuppositions.

4 Experiments 1-2: Presupposition satisfaction and local dependencies

Experiments 1 and 2 asked whether comprehenders are sensitive to locality for presupposition-satisfying dependencies. Since *also* (in a discourse like (2), repeated below as (3)) presupposes only that some other alternative (of the form *Andy bought x*) is true, it might not matter to comprehenders where the content satisfying this presupposition occurs in the preceding discourse.

(3) a. The roommates often go to the farmer's market together.
b. Beth always buys bread.
c. Andy usually buys some celery.
d. His doctor told him he needs to eat more vegetables.
e. Today Andy treated himself to a croissant.
f. He also bought some NECTARINES.

If locality does matter to comprehenders, we can ask whether linear distance matters, or distance measured with respect to hierarchically-structured constituents.[6]

Experiment 1 tested discourses like (3), in which both linear and hierarchical locality predict the final sentence is most easily interpreted as "Andy bought nectarines and a croissant," where "Today Andy treated himself to a croissant" is linearly closest to, and in the smallest discourse constituent containing the trigger *also* (Figure 3

[6] For purposes of this study, I make the simplifying assumptions that sentences are atomic discourse units — i.e. they are not further decomposed, and that they are related to each other by a finite set of discourse connectives, which often but not always correspond to natural language connectives.

gives a hypothetical structured representation of the discourses tested in Experiment 1; tree structures are modeled after Roberts (1996) and Büring (2003)).

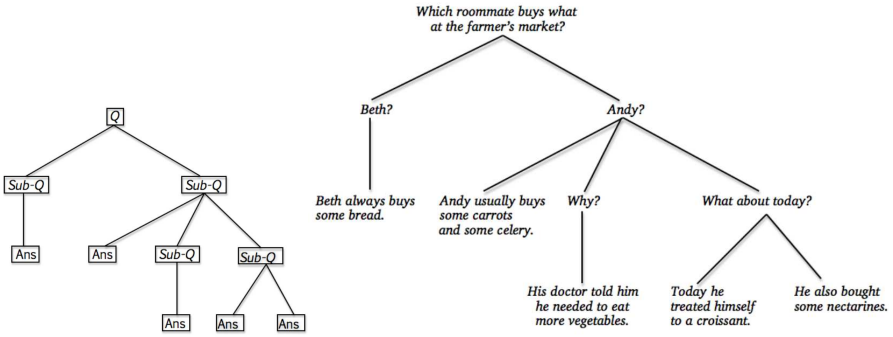

Fig. 3 Discourse tree for Experiment 1.

Experiment 2 tested discourses like (4). Here, unlike the discourses in Experiment 1, what counts as local differs depending on whether distance is measured in terms of linear distance, or in terms of structured representations. If dependencies minimize linear distance, (4-e) should show the same interpretive bias as in (3-f), yielding an interpretation where Andy bought nectarines and croissants. Alternative interpretations of (4-e) should decrease in likelihood as the linear distance spanned by the presupposed content in the preceding discourse increases: for example, the interpretation where Andy bought nectarines, croissants and carrots should be more likely than the interpretation where he bought nectarines, croissants, carrots, and bread. On the other hand, if locality is defined over hierarchical representations (as depicted in Figure 4), the difference in discourse structure between (3) and (4) should yield different interpretations: the closest dominating discourse node is the discourse-initial topic (4-a) for (4-e), but is discourse-medial for (3-f).

(4) a. The roommates went to the farmer's market together.
b. Beth bought some bread.
c. Frank bought some carrots.
d. When his girlfriend is there, she always gets some croissants.
e. Andy also bought some NECTARINES.

4.1 Method

Participants. Twenty native English speakers recruited via Amazon Mechanical Turk participated in Experiment 1; a separate twenty participated in Experiment 2. Compensation depended on how quickly each participant completed the study, averaging a rate of $5.00 per hour.

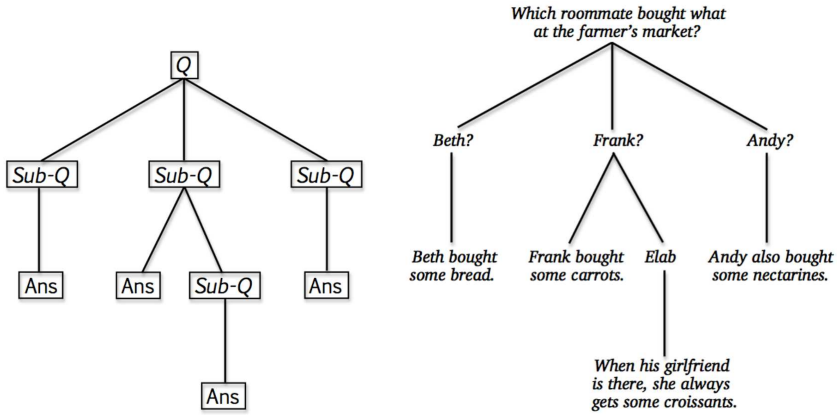

Fig. 4 Discourse tree for Experiment 2.

Materials. The materials for Experiment 1 consisted of ten discourses like (3). In each six-sentence discourse, the first sentence introduced the discourse topic, the second and third sentences introduced two characters, and the remaining three sentences continue being about the second character. The discourse-final sentence features *also*. Participants were instructed to indicate their interpretation of the final sentence in the context of the entire discourse by choosing one of four responses provided. An example of the response types is given in (5), for the discourse in (3).

(5) a. LOCAL INTERPRETATION:
Andy bought some nectarines and a croissant.
b. INTERMEDIATE INTERPRETATION:
Andy bought some nectarines, a croissant, and some celery.
c. GLOBAL INTERPRETATION:
Andy bought some nectarines, a croissant, some celery, and some bread.
d. FALSE INTERPRETATION:
Andy bought a croissant, some celery, and some bread.

The labels for the response types indicate either the distance spanned by the discourse content interpreted as the presupposition (LOCAL, INTERMEDIATE, GLOBAL), or that the interpretation is one that makes the discourse-final sentence false (FALSE).[7] For reference, Table 1 shows how the response types for Experiments 1-2 compare with respect to linear and structured locality.

The materials for Experiment 2 consisted of six discourses with the same form as (4). In each five-sentence discourse, the first sentence introduced the discourse

[7] Note that because the subject of the final sentence in (3) is a pronoun, it is incompatible with the prosodic pattern associated with subject focus. This is not the case for the discourses used in Experiment 2 (4), only responses compatible with direct object focus were available in that experiment.

Table 1 Response types, Experiments 1-3.

Experiment	Response type	Linear locality	Structured locality
1	Local	**local**	**local**
	Intermediate	less local	(non-constituent)
	Global	least local	non-local
	False		
2	Linear local	**local**	(non-constituent)
	Intermediate wrt model	less local	(non-constituent)
	Intermediate wrt mention	less local	(non-constituent)
	Structured local	least local	**local**
3	Linear superset	**local**	(non-constituent)
	Structured superset	non-local	**local**
	Mentioned (false)		
	Mentioned-subset (false)		
	Novel (presupposition failure)		

topic, the second and third sentences introduced two characters, the fourth sentence elaborated on the third sentence, and the final sentence introduced a third character. As in Experiment 1, the discourse-final sentence contained *also*, and participants indicated their interpretation in the context of the discourse by choosing one of four responses. An example of the response types is given in (6), for the discourse in (4).

(6) a. LINEAR LOCAL INTERPRETATION:
Andy bought some nectarines and some croissants.

b. INTERMEDIATE INTERPRETATION (WITH RESPECT TO SITUATION MODEL):
Andy bought some nectarines and some carrots.

c. INTERMEDIATE INTERPRETATION (WITH RESPECT TO DISCOURSE MENTION):
Andy bought some nectarines, some croissants, and some carrots.

d. STRUCTURED LOCAL INTERPRETATION:
Andy bought some nectarines, some carrots, and some bread.

As explained above, the discourses in Experiment 2 were designed to pit linear locality (represented by the LINEAR LOCAL interpretation) against structured locality (represented by the STRUCTURED LOCAL interpretation).

The two INTERMEDIATE interpretations are analogous to the INTERMEDIATE interpretation in Experiment 1: the discourse content included in the presupposition is more than just the most recently mentioned item, but less than the entire set of mentioned items. The interpretations differ in that (6-b) excludes items inconsistent with the situation model of the events described (i.e. the content of the fourth sentence is excluded since it does not describe an actual event), whereas (6-c) includes this content. A bias in favor of interpretation (6-b) relative to (6-c) would suggest that comprehenders interpret presupposed content by referring to a mental model of the situation being described, rather than a discourse representation that only keeps track of whether an item has been mentioned.

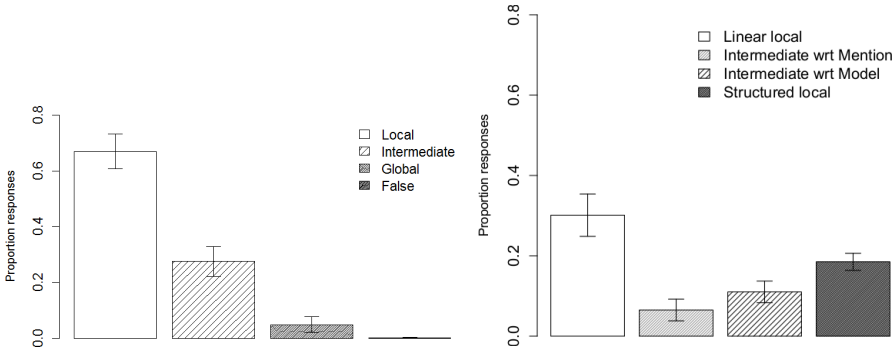

Fig. 5 Experiments 1 (left) and 2 (right) Results: proportion of responses.

4.2 Results and discussion

The results of Experiments 1 and 2 provide support for both a structured view of discourse representations, and sensitivity to linear dependency distance. In both discourse types, comprehenders showed preferences for two interpretations: the material introduced in (i) the most recent sentence in the discourse or (ii) the smallest discourse unit dominating the sentence with the trigger were most likely to be construed as the presupposed content. Both of these interpretations were favored relative to intermediate interpretations, which neither minimized locality in terms of recency, or formed discourse constituents. For Experiment 1 (3), both (i) and (ii) correspond to the interpretation where Andy is understood to have gotten nectarines and a croissant. Figure 5 (left panel) shows the proportions of responses of each type. Pairwise comparisons showed that all pairs except for the Global and False responses differed in the frequency of responses (Table 2).

Table 2 Experiment 1 Pairwise comparisons of response types.

Contrast	χ^2	adjusted p
Local>Intermediate	34.04	< 0.0001
Local>Global	104.34	< 0.0001
Local>False	131.03	< 0.0001
Intermediate>Global	28.45	< 0.0001
Intermediate>False	51.07	< 0.0001
Global>False	8.33	< 0.005

For Experiment 2 (4), minimizing the linear distance spanned by the presupposition dependency corresponds to the interpretation where Andy is understood to have gotten nectarines and croissants. This was the most frequent interpretation, as shown in Figure 5 (right panel). In the second most frequent interpretation, the material introduced in the local discourse unit is construed as the presupposition — here, Andy is understood to have gotten every item mentioned in the prior dis-

course. Pairwise comparisons for the four response types showed that linear-local responses significantly exceeded both intermediate response types, and marginally exceeded structured-local responses. Structured-local responses significantly exceeded intermediate-with-respect-to-mention responses, but not intermediate-with-respect-to-model responses (Table 3). Note that there was a significant number of structured-local responses (t=9.51, $p < 0.0005$), despite the fact that this interpretation violates strict linear locality; by contrast, the global responses in Experiment 1 did not differ significantly from zero (t=1.82, $p = 0.1$).

Table 3 Experiment 2 Pairwise comparisons of response types.

Contrast	χ^2	adjusted p
LinearLocal>IntermediateModel	16.49	< 0.001
LinearLocal>IntermediateMention	21.33	< 0.001
LinearLocal≈StructuredLocal	4.0	$= 0.1$
IntermediateModel≈IntermediateMention	1.64	*n.s.*
IntermediateModel≈StructuredLocal	2.63	*n.s.*
IntermediateMention<StructuredLocal	8.0	< 0.05

It is striking that, as in Experiment 1 (Figure 5), linear-local responses predominated, considering that this interpretation is not even consistent with the situation model comprehenders must construct to interpret the discourse (in the example discourse in (4), no croissants were purchased in the sequence of events described up to the target sentence). This suggests that there is indeed a pressure to minimize linear distance for presupposition dependencies; this recency bias appears to be prioritized over interpretations where the presupposition conforms to the mental model of the situation being described by the discourse. Because the two intermediate response types do not differ reliably in the current data, no definitive conclusions can be drawn about whether presupposition dependencies care about consistency with the situation model. However, to the extent that reasoning is required to suppress readings that are inconsistent with the situation model, the results of Experiment 2 may reflect the intrusion of model-inconsistent readings due to low-level pressure to satisfy a presupposition with the closest available material.

Turning to the structured-local responses, note that the advantage over either of the intermediate responses contrasts strikingly with the pattern of responses in Experiment 1, where intermediate responses reliably exceeded global responses. Since the structured-local responses represent maximal linear distance, this contrast goes against any linear distance minimization constraint. The fact that either the linear-local or structured-local response was chosen more often than either of the intermediate responses suggests that there is both a pressure to minimize linear distance and a bias toward local interpretations that respect discourse units (though in the aggregated data we cannot tell whether both constraints were respected by all participants, or whether participants respected either linear or structured constraint).

Within the relevant local discourse constituent, exhaustive interpretations were preferred to restrictive ones: comprehenders resisted distinguishing among discourse units with the same hierarchical status. That is, they preferred to interpret

sentence (4-e) as meaning that Andy bought all the items mentioned in the discourse, as opposed to e.g. carrots, croissants and nectarines (but not bread). Together, Experiments 1 and 2 provide preliminary offline data suggesting that comprehenders are sensitive to both structure-insensitive recency (minimizing linear distance) and locality in terms of structured discourse representations, even in cases where preserving discourse "constituents" may sacrifice strict linear locality.

5 Experiment 3: Different ways of satisfying the same presupposition

Experiment 3 tracks the timecourse of presupposition resolution in discourses containing *also*. In light of the results of Experiment 2, which implicates both linear and structural locality constraints, the current study asks whether competition between multiple possible interpretations of a presupposition is observed online, and additionally, whether there is evidence of a bias to preserve discourse constituency when constituency-preserving and constituency-violating interpretations are available.

5.1 Method

Participants. Twenty-seven undergraduate students from the University of Rochester participated in Experiment 3. Participants were recruited from introductory Linguistics courses and flyers posted on the university campus, and were paid $7.50 per session. All participants were native speakers of American English, and had normal or corrected-to-normal vision.

Materials and design. Experimental materials consisted of discourses like (7).

(7) a. The roommates went to the farmer's market together.
b. Beth bought some bread.
c. Frank bought some carrots and some apples.
d. Andy also got some NECTARINES.

Each discourse appeared with one of three display types, illustrated in Figure 6. All displays contained at least one subset of mentioned items (8-a)-(8-b) and one set of all discourse-new items (8-c). In addition, the displays included one of the following: (i) a superset of locally-mentioned items in terms of linear distance (*Linear-local display*, (8-d)), (ii) a superset of locally-mentioned items in terms of structured discourse constituency (*Structured-local display*, (8-e)), or (iii) both linearly and structurally defined supersets (*Competition display*; (8-d)-(8-e)).[8]

[8] In order to minimize differences in complexity among display quadrants, each quadrant contained 6-8 objects, regardless of the number of object types present. For example, a *subset* quadrant would

(8) a. subset of mentioned (carrots, apples)
b. subset of mentioned (apples)
c. all novel (nectarines)
d. superset of mentioned (linear) (carrots, apples, nectarines)
e. superset of mentioned (structured) (carrots, apples, bread, nectarines)

Each display contained four 200x200 pixel images located at the corners of the 1024x768 pixel computer screen (with images flush with screen edges).

Participants saw five tokens of each combination of display type (*Linear-local*, *Structured-local* or *Competition*) and discourse type (*Also* or No-*also*), yielding a total of 30 experimental trials. These were interspersed with 68 filler trials design to minimize statistical regularities in the materials. The trials were presented in a random order generated on each run of the experiment. The 98 trials were preceded by four practice trials, none containing *also* or featuring a target item that would appear as a target in an experimental trial.

If comprehenders construct and use hierarchical constituent structures online to constrain presupposition satisfaction, we expect a preference for the structured superset, which respects discourse constituency but violates linear locality, over the linearly defined superset, which minimizes linear distance, but breaks up a discourse constituent. The discourse trees for the interpretations corresponding to the linear and structured superset responses are shown in Figure 7.

Procedure. Each trial began with the participant fixating and clicking on a crosshair in the center of the screen. Participants listened to the discourses over headphones. The display appeared on the computer screen at the onset of the target sentence; there was no preview. Participants were instructed to click on the items that the last-mentioned character got (e.g. "what Andy got"). The trial ended when the participant clicked on a picture in the display. Eye movements were recorded from the onset of the target sentence to the end of the trial, using a head-mounted SR EyeLink II eye-tracking system sampling at 250Hz.

5.2 Results and discussion

Because one of the objectives of Experiment 3 was to observe participants' behavior when more than one viable interpretation is available, I analyze response data and eye movements, and when appropriate, eye movements contingent on response.

Looking first at the Competition display condition, we can see that both of the superset interpretations (linear local or structured local) are possible interpretations of the target sentence: when both interpretations are available in the display, responses are split (albeit unevenly) between the two superset types (Figure 8). In fact, with an online interpretation task, we see a pronounced bias in favor of the structured

have 6-8 objects of the same type (e.g. apples), while a *superset* quadrant might have two of each of four object types (e.g. carrots, apples, bread, nectarines).

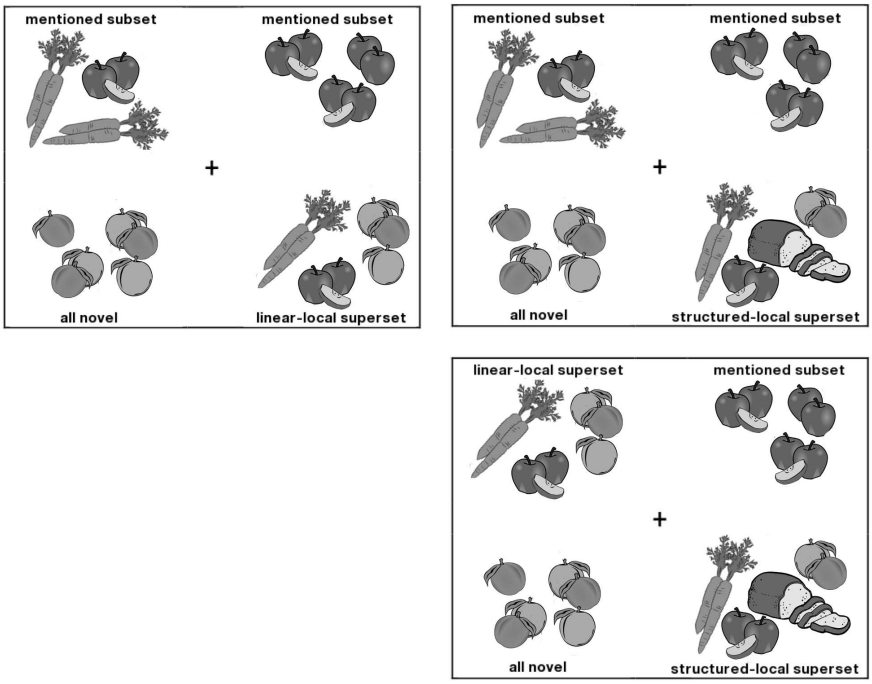

Fig. 6 Experiment 3 display types. Clockwise: *Linear-local, Structured-local* and *Competition displays* (labels for illustration only; illustration — author's own).

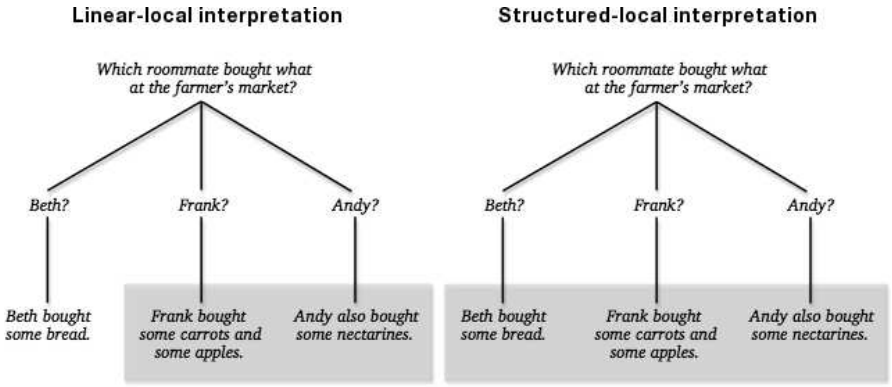

Fig. 7 Discourse tree representation of linear-local and structured-local interpretations.

local interpretation; this contrasts with the offline responses in Experiment 2, where linear local interpretations predominated (Figure 5).

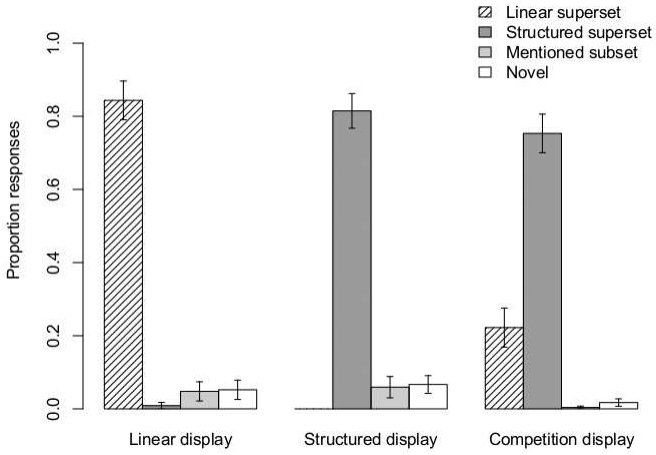

Fig. 8 Experiment 3 Results, Proportions of responses, by display type.

The eye movement data also suggest that the two superset interpretations remain in competition after other options have been ruled out, and in fact well after the offset of the target word. Figure 9 shows the proportion of fixations for the Competition display condition; Figures 10-11 break the data down by response type: Figure 10 represents trials where the participant chose the structured local interpretation, and Figure 11 represents trials where the linear local interpretation was chosen. When the structured interpretation was chosen (Figure 10), fixations to the linear superset reliably exceed fixations to the subset referent in the 400-600 ms window after the target word onset (t=2.81, $p < 0.05$), and this difference persists at least until the 1800-2000 ms window — well after the offset of the target word. When the linear interpretation was chosen (Figure 11), fixations to the structured superset exceed subset fixations beginning in the 1200-1400 ms window (t=2.13, $p < 0.05$); this difference persists until at least 1800-2000 ms after target onset.

Despite the fact that both interpretations appear to be considered online, one interpretation may be preferred over the other — this is suggested by the asymmetry in response types in the Competition display condition: when both interpretations were available in the visual display, participants chose the structured local interpretation more often than they chose the linear local interpretation (Figure 8). This reflects a preference for discourse constituency-preserving interpretation at the expense of minimizing linear dependency distance. To explore timecourse differences between these two interpretations, we will look at the two conditions where only one interpretation was available (*Linear-only* and *Structured-only* display conditions).

Proportions of fixations for *Linear-only* and *Structured-only* display conditions are shown in Figures 12-13. Target fixations from these two conditions (excluding

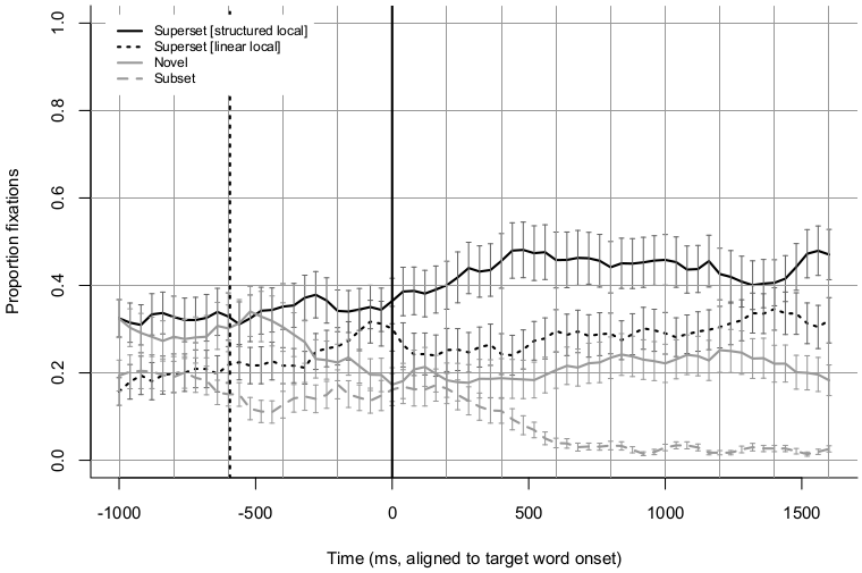

Fig. 9 Experiment 3 Results, Competition display conditions (all response types): Mean proportion of target fixations. (Dotted vertical line=average particle onset; solid vertical line=target word onset.)

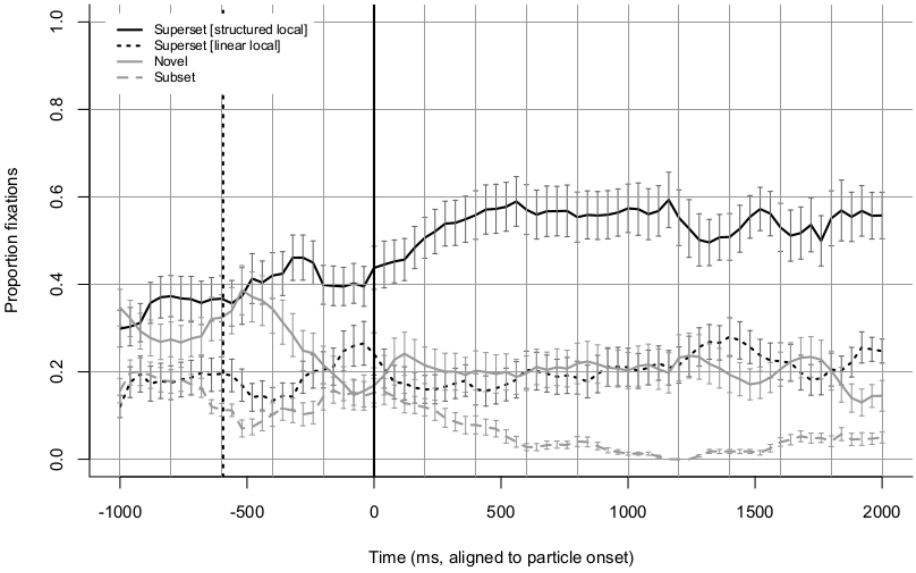

Fig. 10 Experiment 3 Results, Competition display conditions (Structured local responders): Mean proportion of target fixations. (Dotted vertical line=average particle onset; solid vertical line=target word onset.)

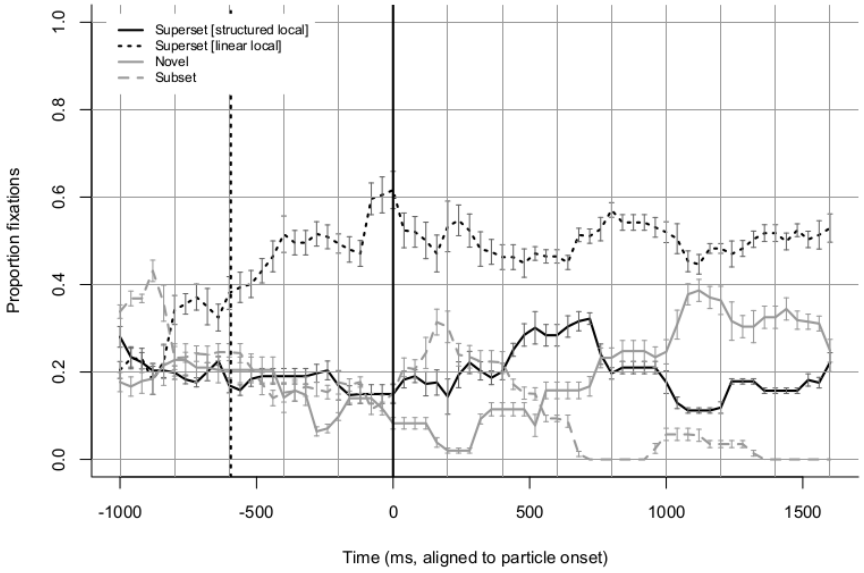

Fig. 11 Experiment 3 Results, Competition display conditions (Linear local responders): Mean proportion of target fixations. (Dotted vertical line=average particle onset; solid vertical line=target word onset.)

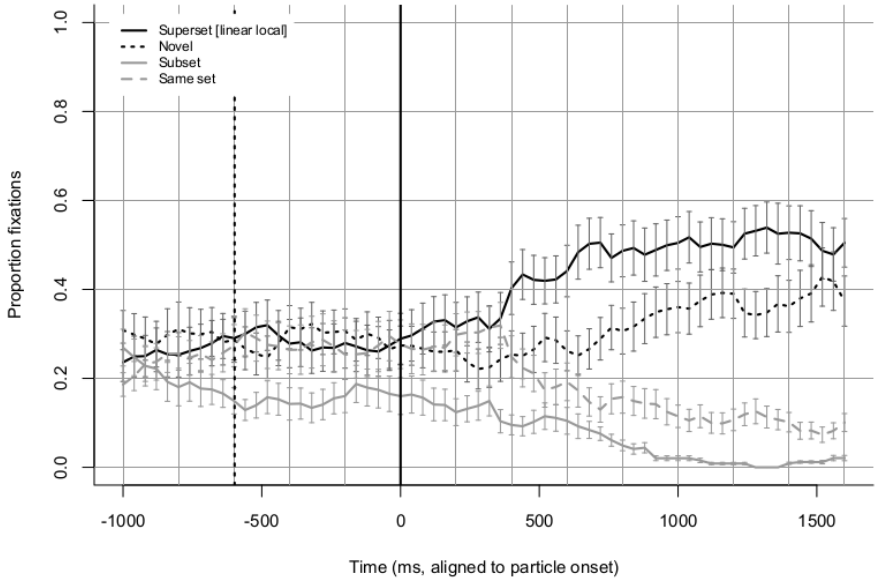

Fig. 12 Experiment 3 Results, Linear-only display conditions: Mean proportion of target fixations. (Dotted vertical line=average particle onset; solid vertical line=target word onset.)

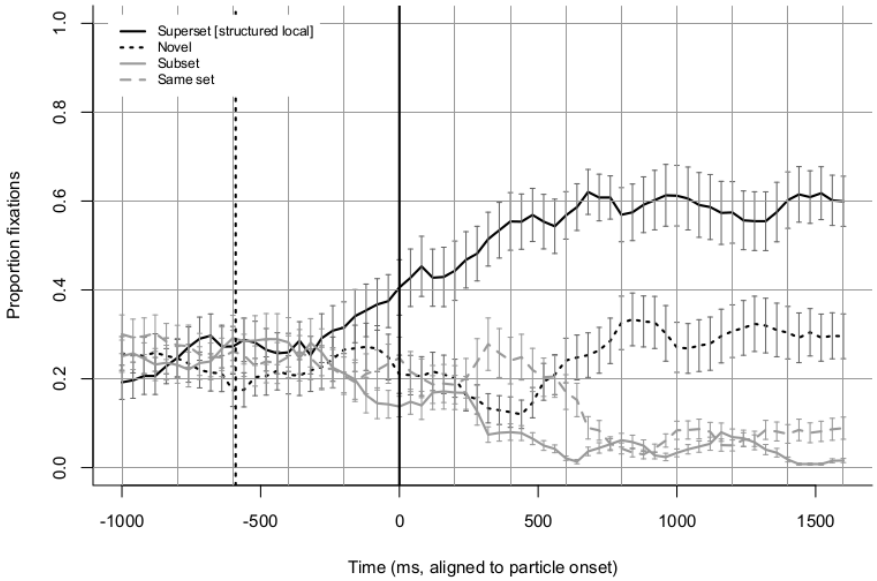

Fig. 13 Experiment 3 Results, Structured-only display conditions: Mean proportion of target fixations. (Dotted vertical line=average particle onset; solid vertical line=target word onset.)

the Competition conditions) were fitted using mixed-effect logistic regression models in three analysis windows, delimited by salient linguistic events in the stimuli: the pre-particle window spans the 500 ms before the onset of "also," the pre-target window starts at the onset of the particle and ends at the onset of the target word, and the post-target window starts at the target word onset and ends 500 ms later. The onsets used to delimit these windows were determined on a trial-by-trial basis. The models predicted fixations to the target referent, and included (1) Display type and (2) Time as fixed effects, and Participant and Item as random effects (Jaeger, 2008; Barr, 2008). Data was unaggregated in all the regression models reported. In addition, we included the state of the previous fixation (on or off the target) as a predictor to deal with the oversampling problem that arises in analyses of Visual World fixation data (Frank et al., 2009).

Analyses began with the full model, which included the interaction between Display type and Time (interactions with State were not included, since there is no clear theoretical reason for State to interact with the variables of interest, and any such interactions would not be interpretable). The State term was left in the model, regardless of significance. All predictors were centered. Redundant terms were removed by eliminating one predictor at a time for all terms correlated with one or more other terms in the model, starting with the highest order term. Model comparison using the likelihood ratio test determined whether the model including the predictor increased the likelihood of the data relative to the model excluding that term. To determine the random effects structure, for each model with a given fixed effects structure, we performed forward step-wise model comparison, starting with

the model containing just random intercepts for Participant and Item, and iteratively adding random effects to the model. Each resulting model was compared to the previous one using the likelihood ratio test. The estimated coefficients for the final models are shown in Tables 4-6 for the initial, early and late analysis windows.

Table 4 Experiment 3 Estimates of fixed effects, initial window.

```
TargetFix ~ DisplayType + Time + State
              + (1+DisplayType|Participant)
              + (1+DisplayType|Item)
```

	Estimate	*SE*	z	*p*
Intercept	-7.15	0.76	-9.40	< 0.0001
StructuredLocalDisplay	0.05	0.24	0.20	*n.s.*
Time	-1.13	0.83	-1.37	*n.s.*
State	11.93	0.28	42.69	< 0.0001

Table 5 Experiment 3 Estimates of fixed effects, early window.

```
TargetFix ~ DisplayType + Time + State + DisplayType:Time
              + (1+DisplayType|Participant)
              + (1+DisplayType|Item)
```

	Estimate	*SE*	z	*p*
Intercept	-6.67	0.37	17.85	< 0.0001
StructuredLocalDisplay	0.88	0.46	1.94	0.05
Time	-0.36	0.97	-0.37	*n.s.*
State	12.01	0.23	51.64	< 0.0001
StructuredDisplay:Time	1.99	1.29	1.55	0.12

Table 6 Experiment 3 Estimates of fixed effects, late window.

```
TargetFix ~ DisplayType + Time + State
              + (1+DisplayType|Participant)
              + (1+DisplayType|Item)
```

	Estimate	*SE*	z	*p*
Intercept	-5.75	0.24	-24.11	< 0.0001
StructuredLocalDisplay	0.14	0.20	0.69	*n.s.*
Time	0.08	0.68	-0.11	*n.s.*
State	11.36	0.21	55.07	< 0.0001

As can be seen both in the models in Tables 4 and 6, and by comparing the fixation plots in Figures 12 and 13, there is no effect of Display type on target fixations in the initial (500 ms preceding particle onset) or late (500 ms following target onset) analysis windows. However, in the early window (particle onset to target onset), there is a main effect of Display type: participants were more likely to fixate the

eventual superset target when the display provided only a structured-local interpretation, relative to displays providing only a linear-local interpretation (Table 5). The model also includes a positive Display type by Time which does not reach significance. This asymmetry between two interpretations — both of which are possible — is also reflected in response times. Participants were slower to respond by mouse-click when they were choosing a linear interpretation than when they were choosing a structured interpretation (Figure 14).[9]

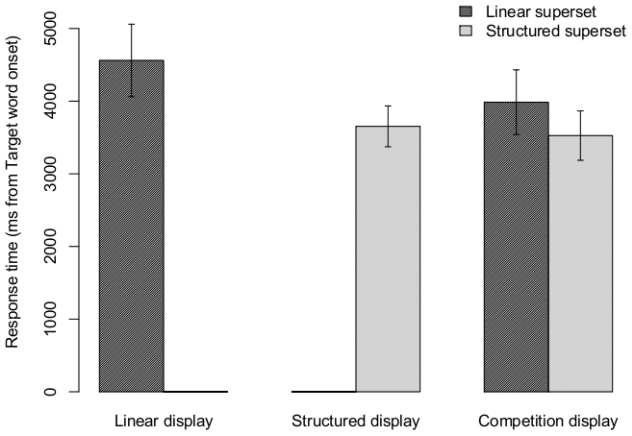

Fig. 14 Experiment 3 Results, Response times (mouse click), by response type and display type.

These results suggest that the linear interpretation is dispreferred relative to the structured interpretation, even when it is the only display item that satisfies the presupposition of *also*. Figure 15 shows a structured representation of (7-d). Since the linear and structured superset interpretations force the comprehender to interpret the immediately preceding sentence and the entire discourse, respectively, as the presupposed material, this amounts to a preference for satisfying the presupposition at the level of the smallest discourse unit containing the presupposition trigger, regardless of linear distance.

As in Experiments 1 and 2, these data show that comprehenders prefer to minimize dependency length in interpreting presupposed content. The timecourse information from Experiment 3 suggests that, in online discourse processing, interpretations that rely on structured representations may be more available than ones which rely on linear precedence alone.

[9] The rightmost bars in Figure 14 also show a numerical advantage for response times in the Competition condition when the structured interpretation was chosen, compared to when the linear interpretation was chosen.

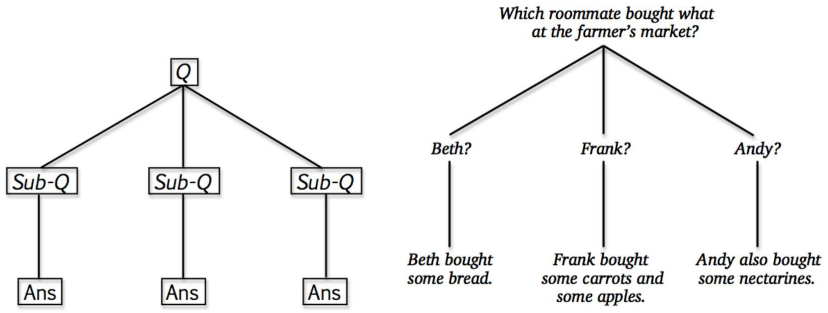

Fig. 15 Discourse tree for Experiment 3.

6 Discussion

Together, the results of Experiments 1-3 suggest that, as in other domains of processing, comprehenders favor local dependencies. However, we see evidence both for locality defined by linear distance, and locality defined over hierarchical discourse structures. I leave for further investigation the question of how these constraints play off against each other given factors such as whether multiple viable interpretations are available, and whether the response measure is online.

Although the preceding discussion has characterized the salient interpretations as differing in terms of whether they are based on linear or structured discourse representations, the current data still allow for explanations that do not make reference to structured discourse representations. I outline one such account here, in concluding the discussion of Experiment 3.

According to a view of sentence processing as cue-based memory retrieval Lewis and Vasishth (2005); Lewis et al. (2006), each word triggers dependency formation with preceding material in the sentence. This prior material is retrieved on the basis of associative cues, where retrieval cues of the current word partially match the features of the material to be retrieved. For instance, a verb in a relative clause will trigger the retrieval of its argument, forming a syntactic dependency between the elements underlined in (9).

(9) Sameer bought the book that Justin recommended.

Such an account might be able to account for the interpretive biases observed in Experiments 1-3, if extended to discourse-level dependencies. Recall that there was evidence for both an interpretation that satisfied the presupposition of *also* using the closest available material, as well as an interpretation that satisfied the presupposition with all the material in the prior discourse, going back to a local discourse topic. While the above discussion characterized the latter interpretation as a structure-sensitive one, it might also be favored by associative cue-based retrieval. The example discourse from Experiment 2 is repeated below as (10); the material matching

the focus in grammatical function, syntactic category, and conceptual features — candidates for retrieval by a cue-based retrieval mechanism — is underlined.

(10) a. The roommates went to the farmer's market together.
b. Beth bought some bread.
c. Frank bought some carrots.
d. When his girlfriend is there, she always gets some croissants.
e. Andy also bought some NECTARINES.

The structured-local interpretation from Experiments 1-3 represents one where all material (in the search window) is retrieved based on featural similarity with the focused element.

The linear-local interpretation, on the other hand, is the interpretation that retrieves the minimal material from the prior discourse that is a partial featural match with the focus — if search proceeds backward from the focused element, this will be material in the immediately prior sentence (in (10)). The data from Experiments 1-3 might be interpreted as an interplay between these two pressures — to maximize feature match with the dependency trigger, and to retrieve only as little as necessary to form the dependency — rather than a difference in the discourse representations underlying the salient interpretations. In future research, studying the interpretations available in more complex discourses will help test the predictions of these contrasting explanations where they diverge.

I will close with a speculative note about cross-linguistic variation and generality. What aspects of language and how we process language should we expect to be more or less stable, irrespective of the properties of particular languages? And on the other hand, what should we expect to vary as a function of language-specific properties? The stable things will likely be things that we may think of as linguistic because they are inextricably tied to linguistic competence and performance, but which are not in and of themselves linguistic — for instance, language is affected by considerations like likelihood of communicative success, communicative efficiency (Aylett and Turk, 2004; Levy and Jaeger, 2007; Jaeger, 2010); we plan and interpret language based on our estimations of common ground and perspective information (Clark, 1992, 1996; Hanna et al., 2003), and the goals and intentions of our interlocutors (Grosz and Sidner, 1986; Ballard and Hayhoe, 2009); inferences we make about the knowledge states of our interlocutors then combines with features of the utterance context to determine what is salient (Brown-Schmidt et al., 2008; Heller et al., 2008; Arnold, 2010). Modulo cultural differences that may influence the shape these basic processes occur, the influence they exert on language use should be stable.

On the other hand, what *should* we expect to vary as a function of language-specific properties? Take locality preferences as an example. The characterization of dependency length minimization in e.g. Hawkins (1994) suggests that it should be considered to have broad, non-language-specific application. But it has been shown that certain languages show a locality bias while others show the opposite — what looks like an anti-locality bias (Vasishth and Lewis, 2006). If we assume there is a very general bias toward minimizing working memory load which applies to a

wide range of cognitive processes including language processing, then language-specific properties like head directionality might systematically predict differences in local/anti-local dependency preference — a proposal along these lines is made by Vasishth and Lewis (2006). Another point to keep in mind moving forward is precisely how general processes, such as those underlying intention recognition or the integration of sensory inputs, makes contact with aspects of language use and competence that seem truly language-internal, such as mental representations of sentence or discourse structure. Possibly, some of the apparently truly language-internal elements will turn out to receive more parsimonious explanations as instances of broader phenomena; if so, it will be even more interesting to see what things we cannot explain by invoking general-purpose cognitive processes. The research presented here represents an effort to understand some empirical observations about a linguistic phenomenon from this perspective.

References

J. E. Arnold. How speakers refer: the role of accessibility. *Language and Linguistics Compass*, 4(4):187 – 203, 2010.

J. E. Arnold and Z. M. Griffin. The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, 56: 521–536, 2007.

N. Asher and A. Lascarides. *Logics of Conversation*. Cambridge University Press, 2003.

J. D. Atlas. Topic/comment, presupposition, logical form and focus stress implicatures: The case of focal particles *only* and *also*. *Journal of Semantics*, 8:127–147, 1991.

M. Aylett and A. Turk. The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence and duration in spontaneous speech. *Language and Speech*, 47(1):31 – 56, 2004.

D. Ballard and M. Hayhoe. Modeling the role of task in the control of gaze. *Visual Cognition*, 17:1185 – 1204, 2009.

D. J. Barr. Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59:457 – 474, 2008.

R. de Beaugrande and B. N. Colby. Narrative models of action and interaction. *Cognitive Science*, 3:43–66, 1979.

K. Bicknell and H. Rohde. Dynamic integration of pragmatic expectations and real-world event knowledge in syntactic ambiguity resolution. In *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, pages 1216–1221, 2009.

S. Brown-Schmidt, C. Gunlogson, and M. K. Tanenhaus. Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition*, 107:1122 – 1134, 2008.

D. Büring. On d-trees, beans, and b-accents. *Linguistics and Philosophy*, 26:511–545, 2003.

H. H. Clark. *Arenas of Language Use*. University of Chicago Press, Chicago, 1992.

H. H. Clark. *Using language*. Cambridge University Press, Cambridge, 1996.

P. Cohen and R. Perrault. Elements of a plan-based theory of speech acts. *Cognitive Science*, 3:177–212, 1979.

R. Cooper and J. Ginzburg. *Using dependent record types in clarification ellipsis*, volume Proceedings of the sixth workshop on the semantics and pragmatics of dialogue, pages 45–52. Edinburgh University, 2002.

D. Farkas and K. Bruce. On reacting to assertions and polar questions. *Journal of Semantics*, 2009.

A. Frank, A. P. Salverda, T. F. Jaeger, and M. K. Tanenhaus. Analyzing multinomial time-series with "state" dependencies. Poster presented at the 22nd CUNY Sentence Processing Conference, Davis, CA, 2009.

L. Frazier and C. Clifton Jr. *Construal*. MIT Press, Cambridge, Mass., 1996.

L. Frazier and C. Clifton Jr. Ellipsis and discourse coherence. *Linguistics and Philosophy*, 29(3):316–346, 2006.

J. P. Gee and F. Grosjean. Empirical evidence for narrative structure. *Cognitive Science*, 8:59–85, 1984.

E. Gibson. Linguistic complexity: locality of syntactic dependencies. *Cognition*, 68:1 – 76, 1998.

E. Gibson. The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz, and W. O'Neill, editors, *Image, language, brain*, pages 95 – 126. MIT Press, Cambridge, MA, 2000.

J. Ginzburg. *The Interactive Stance: Meaning for Conversation*. Oxford University Press, 2012.

J. Ginzburg and I. A. Sag. *Interrogative Investigations: the form, meaning and use of English interrogative constructions*. CSLI, 2000.

A. M. Glenberg, A. C. Wilkinson, and W. Epstein. The illusion of knowing: Failure in the self-assessment of comprehension. *Memory and Cognition*, 10:597 – 602, 1982.

A. C. Graesser and C. L. McMahen. The influence of connectives on sentence comprehension. *Journal of Memory and Language*, 33:128–147, 1993.

A. C. Graesser and T. Singer, M. ad Trabasso. Constructing inferences during narrative text comprehension. *Psychological Review*, 101:371–395, 1994.

J. Groenendijk and M. Stokhof. *Studies in the Semantics of Questions and the Pragmatics of Answers*. PhD thesis, University of Amsterdam, The Netherlands, 1984.

B. Grosz and C. Sidner. Attention, intention, and the structure of discourse. *Computational Linguistics*, 12(3):175 – 204, 1986.

B. Grosz, A. Joshi, and S. Weinstein. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 2(21):203 – 225, 1995.

J. E. Hanna, M. K. Tanenhaus, and J. C. Trueswell. The eects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, 49:43 – 61, 2003.

J. A. Hawkins. *A performance theory of order and constituency*. Cambridge University Press, Cambridge, UK, 1994.

J. A. Hawkins. *Efficiency and complexity in grammars*. Oxford University Press, Oxford, UK, 2004.

P. Hayes. *Some association-based techniques for lexical disambiguation by machine*. PhD thesis, University of Rochester, 1977. Published as Technical Report No. 25, Department of Computer Science.

I. Heim. *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, Amherst, 1982.

D. Heller, D. Grodner, and M. K. Tanenhaus. The role of perspective in identifying domains of reference. *Cognition*, 108(831 – 836):831 – 836, 2008.

B. Hemforth, L. Konieczny, and C. Scheepers. Syntactic attachment and anaphor resolution: The two sides of relative clause attachment. In M. Crocker, M. Pickering, and C. Clifton Jr., editors, *Architectures and mechanisms for language processing*, pages 259 – 282. Cambridge University Press, Cambridge, 2000.

J. Hintikka. The semantics of questions and the questions of semantics. *Acta Philosophica Fennica*, 28, 1976.

J. R. Hobbs. Coherence and coreference. *Cogntive Science*, 3:67–90, 1979.

L. Horn. A presuppositional analysis of *only* and *even*. In *Proceedings of CLS 5*, pages 97–108, Columbus, OH, 1969.

E. H. Hovy. Automated discourse generation using discourse structure relations. *Artificial Intelligence*, 63:341–385, 1993.

T. F. Jaeger. Categorical Data Analysis: Away from ANOVAs (transformation or not) and towards Logit Mixed Models. *Journal of Memory and Language*, 59: 434 – 446, 2008.

T. F. Jaeger. Redundancy and reduction: Speakers manage syntactic information density. *Cognitive Psychology*, 61(1):23 – 62, 2010.

H. Kamp and U. Reyle. *From Discourse to Logic*. Kluwer, Dordrecht, 1993.

L. Karttunen and S. Peters. Conventional implicature. In C.-K. Oh and D. A. Dinneen, editors, *Syntax and Semantics, Volume 11: Presupposition*, pages 1–56. Academic Press, 1979.

A. Kehler. *Interpreting Cohesive Forms inthe Context of Discourse Inference*. PhD thesis, Harvard University, 1995.

A. Kehler. Coherence and the resolution of ellipsis. *Linguistics and Philosophy*, 23 (6):533 – 575, 2000.

A. Kehler. *Coherence, Reference, and the Theory of Grammar*. CSLI Publications, 2002.

A. Kehler, L. Kertz, H. Rohde, and J. L. Elman. Coherence and coreference revisited. *Journal of Semantics: Special Issue on Processing Meaning*, 25:1–44, 2008.

C. S. Kim. *Generating alternatives: Interpreting focus in discourse*. PhD thesis, University of Rochester, 2012.

C. S. Kim and J. T. Runner. Discourse structure and parallelism in VP ellipsis. In *Proceedings of Semantics and Linguistic Theory*, Columbus, OH, 2009.

J. King and M. A. Just. Individual differences in syntactic processing: the role of working memory. *Journal of Memory and Language*, 30:580 – 602, 1991.

W. Kintsch and T. van Dijk. Towards a model of text comprehension and production. *Psychological Review*, 85:363–394, 1978.

A. Knott and T. J. Sanders. The classification of coherence relations and their linguistic markers: An exploration of two languages. *Journal of Pragmatics*, 30: 135–175, 1998.

J. A. Levin and J. A. Moore. Dialogue-games: Metacommunication structures for natural language interaction. *Cognitive Science*, 1(4):395–420, 1977.

R. Levy and T. F. Jaeger. Speakers optimize information density through syntactic reduction. In *Proceedings of the Twentieth Annual Conference on Neural Information Processing Systems*, 2007.

D. K. Lewis. Scorekeeping in a language game. *Journal of Philosophical Logic*, 8: 339–359, 1979.

R. L. Lewis and S. Vasishth. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29:375–419, 2005.

R. L. Lewis, S. Vasishth, and J. A. Van Dyke. Computational principles of working memory in sentence comprehension. *Trends in Cognitive Science*, 10(10):447–454, 2006.

D. J. Litman and J. F. Allen. A plan recognition model for subdialogues in conversation. *Cognitive Science*, 11:163–200, 1987.

W. C. Mann and S. A. Thompson. Rhetorical structure theory: Towards a theory of text organization. *Text*, 8:243–281, 1988.

K. R. Mc-Keown. *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text*. Cambridge University Press, Cambridge, 1985.

K. K. Millis and M. A. Just. The influence of connectives on sentence comprehension. *Journal of Memory and Language*, 33:128–147, 1994.

D. G. Morrow. Prominent characters and events organize narrative understanding. *Journal of Memory and Language*, 24:304–319, 1985.

K. Moulton. Small antecedents: Syntax or pragmatics? In E. Elfner and M. Walkow, editors, *Proceedings of NELS 37*. GLSA Publications, 2006.

G. L. Murphy and A. M. Shapiro. Forgetting of verbatim information in discourse. *Memory and Cognition*, 22:85–94, 1994.

J. Otero and W. Kintsch. Failures to detect contradictions in a text: What readers believe versus what they read. *Psychological Science*, 3:229–235, 1992.

L. Polanyi and M. van den Berg. Discourse structure and discourse contexts. In P. Dekker and M. Stokhof, editors, *Proceedings of the Tenth Amsterdam Colloquium*. ILLC/Department of Philosophy, University of Amsterdam, 1996.

C. Roberts. Towards an integrated formal theory of pragmatics. In Jae Hak Yoon and Andreas Kathol, editors, *OSU Working Papers in Linguistics*, volume 49. The OSU Department of Linguistics, 1996.

H. Rohde. *Coherence-Driven Effects in Sentence and Discourse Processing*. PhD thesis, University of California San Diego, 2008.

H. Rohde, A. Kehler, and J. L. Elman. Pronoun interpretation as a side effect of discourse coherence. In *Proceedings of the 29th Annual Conference of the Cognitive Science Society*, pages 617–622, 2007.

H. Rohde, R. Levy, and A. Kehler. Anticipating explanations in relative clause processing. *Cognition*, 118:339–358, 2011.

D. Roland, C. O'Meara, H. Yun, and G. Mauner. Discourse and object relative clauses: The effect of topic versus mention. Paper presented at the CUNY Conference on Human Sentence Processing, University of North Carolina at Chapel Hill, 2008.

M. Rooth. *Association with Focus*. PhD thesis, University of Massachusetts at Amherst, 1985.

T. J. Sanders, W. P. M. Spooren, and L. G. M. Noordman. Toward a taxonomy of coherence relations. *Discourse Processes*, 15:1 – 35, 1992.

T. J. Sanders, W. P. M. Spooren, and L. G. M. Noordman. Coherence relations in a cognitive theory of discourse representation. *Cognitive Linguistics*, 4(2):93–133, 1993.

R. C. Schank, G. C. Collins, E. Davis, P. N. Johnson, S. Lytinen, and B. J. Reiser. What's the point? *Cognitive Science*, 6(3):255–275, 1982.

F. Schwarz. Processing presupposed content. *Journal of Semantics*, 24(4):373–416, 2007.

J. Simner and M. Pickering. Planning causes and consequences in discourse. *Journal of Memory and Language*, 52:226–239, 2005.

R. Stevenson, A. Knott, J. Oberlander, and S. McDonald. Interpreting pronouns and connectives: interactions among focusing, thematic roles and coherence relations. *Language and Cognitive Processes*, 15:225–262, 2000.

M. K. Tanenhaus, M. J. Spivey-Knowlton, K. M. Eberhard, and J. C. Sedivy. Intergration of visual and linguistic information in spoken language comprehension. *Science*, 268:1632–1634, 1995.

D. Temperley. Minimization of dependency length in written english. *Cognition*, 105:300 – 333, 2007.

M. J. Traxler, M. J. Pickering, and C. Clifton Jr. Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, 39:558–592, 1998.

J. J. van Berkum, C. M. Brown, and P. Hagoort. Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, 41:147–182, 1999.

R. A. van der Sandt. Presupposition projection as anaphora resolution. *Journal of Semantics*, 9(4):333–377, 1992.

J. van Kuppevelt. Discourse structure, topicality and questioning. *Journal of Linguistics*, 1995.

S. Vasishth and R. Lewis. Argumrnt-head distance and processing complexity: Explaining both locality and antilocality effects. *Language*, 82:767 – 794, 2006.

B. Webber and A. Joshi. Anchoring a lexicalized tree-adjoining grammar for discourse. In *Proceedings of the ACL/COLING Workshop on Discourse Relations and Discourse Markers*, Montreal, 1998.

F. Wolf and E. Gibson. Representing discourse coherence: A corpus-based study. *Computational Linguistics*, 31(2):249–288, June 2005. ISSN 0891-2017. doi: 10.1162/0891201054223977. URL http://dx.doi.org/10.1162/0891201054223977.