VISA - Corpus Annotation with OWL *

Stephanie Becker and **Thomas Kleinbauer** and **Stephan Lesch**
DFKI Gmbh
Stuhlsatzenhausweg 3
66123 Saarbrücken, Germany
<firstname.lastname>@dfki.de

Abstract

We present VISA, a graphical annotation tool for OWL-based annotation schemes with a focus on generality and usability.

1 Introduction

The W3C standard OWL was originally designed as an ontology language for the semantic web, but it is progressively finding its way into various other fields of application. Annotated (linguistic) corpora, on the other hand, still often rely on their own specific data storage formats, although newer developments show a trend towards the use of XML (Carletta et al., 2005).

We believe that OWL is a suitable format for future corpora and annotations thereof, as it provides a semantically potent language based on a simple and open format. The main advantage is that further processing of corpus data can make use of automatic inference mechanisms, working only on one underlying formalism for all annotations. Existing annotation schemes can easily be expressed in OWL; annotation then becomes a process of assigning instances of ontology classes to corpus segments.

A number of tools specialized for different kind of annotations exist, as well as programs for working with OWL data. However, the number of tools for annotating OWL ontologies is rather small. One way to build such tools is to combine existing software for annotation and for OWL – a procedure taken for instance by (Bontcheva et al., 2004) or (Lauer et al., 2005) which both integrate the Protégé [1] editor for OWL into their own annotation framework.

But this approach suffers from the fact that Protégé was not originally designed for annotation work. Ontology instances, for example, are displayed as a flat list which makes it difficult for the annotator to discern which corpus segment was annotated with which instances. Relations between instances are displayed in a similar fashion. Furthermore, we found that Protégé reactivity decreases notably with increasing ontology size.

Hence, although a tool that combines existing programs is commendable in principal, practical application may prove very difficult under certain circumstances in which the user might prefer a tool tailored specifically to annotation with OWL. Furthermore, these observations illustrate the importance of good usability for annotation tools.

2 The VISA Annotation Tool

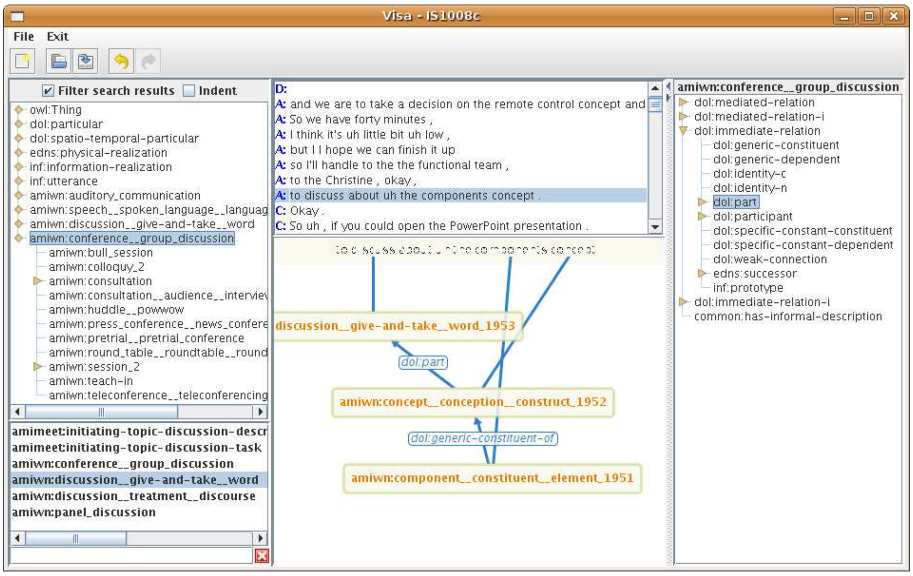
Based on the analysis of deficiencies of existing annotation tools we derived a first requirements specification for a new tool which was followed by the development of a prototype. The further development process has been accompanied by further theoretical considerations with respect to the possible extension of the requirements specification. Moreover we have conducted practical evaluations in form of repeated testing and the prototype has continuously been adapted according to the extended requirements specification.

The following screenshot displays the VISA tool. On the left hand side the classes of the ontology are displayed with their hierarchical relationships, on the right hand side the relation hierarchy of the ontology is shown. In the middle of the window an annotation panel and the text segments that are to be annotated are displayed.

To create a class instance during the annotation process, the corresponding class is selected in the

* The research presented here is funded by the EU under the grant FP6-506811 (AMI).

[1] http://protege.stanford.edu


Visa - IS1008c
File
Exit
☑ Filter search results
□ Indent
owl:Thing
dol:particular
dol:spatio-temporal-particular
edns:physical-realization
inf:information-realization
inf:utterance
amiwn:auditory_communication
amiwn:speech__spoken_language__languag
amiwn:discussion__give-and-take__word
amiwn:conference__group_discussion
amiwn:bull_session
amiwn:colloquy_2
amiwn:consultation
amiwn:consultation__audience__intervie
amiwn:huddle__powwow
amiwn:press_conference__news_confere
amiwn:pretrial__pretrial_conference
amiwn:round_table__roundtable__round
amiwn:session_2
amiwn:teach-in
amiwn:teleconference__teleconferencing
amimeet:initiating-topic-discussion-descr
amimeet:initiating-topic-discussion-task
amiwn:conference__group_discussion
amiwn:discussion__give-and-take__word
amiwn:discussion__treatment__discourse
amiwn:panel_discussion
discussion__give-and-take__word_1953
dol:part
amiwn:concept__conception__construct_1952
dol:generic-constituent-of
amiwn:component__constituent__element_1951
amiwn:conference__group_discussion
dol:mediated-relation
dol:mediated-relation-i
dol:immediate-relation
dol:generic-constituent
dol:generic-dependent
dol:identity-c
dol:identity-n
dol:part
dol:participant
dol:specific-constant-constituent
dol:specific-constant-dependent
dol:weak-connection
edns:successor
inf:prototype
dol:immediate-relation-i
common:has-informal-description

D:
A: and we are to take a decision on the remote control concept and
A: So we have forty minutes ,
A: I think it's uh little bit uh low ,
A: but I I hope we can finish it up
A: so I'll handle to the the functional team ,
A: to the Christine , okay ,
A: to discuss about uh the components concept .
C: Okay .
C: So uh , if you could open the PowerPoint presentation .

hierarchy. An instance of the selected class is then created on the annotation panel by drag and drop.

Class instances can be connected with one or several words of the current text segment by dragging from the instances to the words. Relations between instances can be annotated by selecting a relation from the relation hierarchy and dragging from the instance of the corresponding domain class to the instance of the range class.

The graphical instances are arranged automatically on the annotation panel, thus the annotator does not need to take care of the graphical layout of the annotation. To facilitate navigation in the ontology, keyword search functions are available.

VISA is capable of dealing with large-sized ontologies without slowing down the annotation process. One of the ontologies we tested VISA with , e. g., contains more than 60.000 concepts.

VISA is based on NXT (Carletta et al., 2003) which supports the development of corpus tools through the provision of an open source Java API. However, through its modular architecture, VISA allows the integration of other data formats as well.

3 Conclusion and Future Work

We developed a tool for the annotation of text segments with OWL-based ontologies, focussing on a rich feature set an good usability. VISA can deal with large-sized ontologies without slowing down the annotation process.

VISA requires that the text to be annotated is pre-segmented. Furthermore an already existing ontology is required. As our primary concern is to provide an appropriate tool for annotation, VISA does not provide functions for creating or editing ontologies, nor for segmenting or editing of the corpus.

Currently, VISA should still be considered as a prototype. Several features are planned to be added, particularly with regard to the further facilitation of the annotation process, but also features like a reasoning function in order to prohibit inconsistent annotations.

References

K. Bontcheva, V. Tablan, D. Maynard, and H. Cunningham. 2004. Evolving GATE to Meet New Challenges in Language Engineering. *Natural Language Engineering*, 10(3/4):349–373.

Jean Carletta, Stefan Evert, Ulrich Heid, Jonathan Kilgour, Judy Robertson, and Holger Voormann. 2003. The NITE XML Toolkit: flexible annotation for multi-modal language data. *Behavior Research Methods, Instruments, and Computers, special issue on Measuring Behavior*, 35(3):353–363.

Jean Carletta, Simone Ashby, Sebastien Bourban, Mike Flynn, Mael Guillemot, Thomas Hain, Jaroslav Kadlec, Vasilis Karaiskos, Wessel Kraaij, Melissa Kronenthal, Guillaume Lathoud, Mike Lincoln, Agnes Lisowska, Iain McCowan, Wilfried Post, Dennis Reidsma, and Pierre Wellner. 2005. The ami meeting corpus: A pre-annoncement. In *Proceedings of MLMI 2005*.

Christoph Lauer, Jochen Frey, Benjamin Lang, Jan Alexandersson, Tilman Becker, Thomas Kleinbauer, and Harald Lochert. 2005. Amigram–a general-purpose tool for multimodal corpus annotation. In *Proceedings of MLMI 2005*, Royal College of Physicians, Edinburgh, UK, 11-13 July.

Browsing Meetings: Automatic Understanding, Presentation and Feedback for Multi-Party Conversations*

Patrick Ehlen, Stéphane Laidebeure, John Niekrasz,
Matthew Purver, John Dowding and **Stanley Peters**
Center for the Study of Language and Information
Stanford University
Stanford, CA 94305, USA
{ehlen, laidebeu, niekrasz, mpurver, jdowding, peters}
@csli.stanford.edu

Abstract

We present a system for extracting useful information from multi-party meetings and presenting the results to users via a browser. Users can view automatically extracted discussion topics and action items, initially seeing high-level descriptions, but with the ability to click through to meeting audio and video. Users can also add value: new topics can be defined and searched for, and action items can be edited or corrected, deleted or confirmed. These feedback actions are used as implicit supervision by the understanding agents, retraining classifier models for improved or user-tailored performance.

1 Introduction

Research on multi-party dialogue in meetings has yielded many *meeting browser* tools geared toward providing visual summaries of multimodal data collected from meetings (Tucker and Whittaker, 2005). Why create another? Existing tools focus on facilitating manual annotation and analysis of abstracted knowledge, or on assisting the meeting process by allowing users to conveniently (but manually) add relevant information online.

Because our aim in the CALO Meeting Assistant project is to automatically extract useful information such as the topics and action items discussed during meetings, our meeting browser has a different goal. Not only do we need an end-user-focused interface for users to browse the audio, video, notes, transcripts, and artefacts of meetings, we also need a browser that presents automatically extracted information from our algorithms in a convenient and intuitive manner. And that browser should allow – even compel – users to modify or correct information when automated recognition falls short of the mark.

2 Automatic Understanding

User studies (Banerjee et al., 2005) show that amongst the most requested pieces of information from a meeting are the *topics* discussed and *action items* established.

Action Item Identification. Our understanding suite therefore includes an agent for action item identification – see (Purver et al., 2006). We exploit a shallow notion of discourse structure, by using a hierarchical combination of supervised classifiers. Each sub-classifier is trained to detect a class of utterance which makes a particular discourse contribution to establishing an action item: proposal or description of the related *task*; discussion of the *timeframe* involved; assignment of the responsible party or *owner*; and *agreement* by the relevant people. An overall decision is then made based on local clusters of multiple discourse contributions, and the properties of the hypothesized action item are taken from contributing utterances (the surface strings, semantic content or speaker/addressee identity). Multiple alternative hypotheses about action items and their properties are provided and scored using the individual sub-classifier confidences.

Topic Identification. Another agent splits meetings into topically coherent segments, providing models of the associated topics using vector space

*This work was supported by DARPA grant NBCH-D-03-0010. The content of the information in this publication does not necessarily reflect the position or the policy of the US Government, and no official endorsement should be inferred.

models. Topics are extracted as probability distributions over words, learnt over multiple meetings and stored in a central topic pool; they can then be used for audio/video browsing (labelled via the top most distinctive words) or to interpret a user keyword or sentence search query (by finding the weighted mixture of learnt topics which best match the words of the query).

3 User Interface

Agents that generate multiple hypotheses fare better with feedback from users about which hypotheses sound reasonable, but getting that feedback isn't always easy. A meeting browser is the ideal place to solicit feedback from end-users about what happened during a meeting. Our browser interface exploits the *transparency of uncertainty* principle, which counts on people's tendency to feel compelled to correct errors when those errors are (a) glaringly evident, and (b) correctable in a facile and obvious way.

A user can view action items detected from the meeting in the browser and drag them to a bin that adds the items to the user's to-do list. For the properties of action items – such as their descriptions, owners, and timeframes – the background colors of hypotheses are tied to their sub-classifier confidence scores, so less certain hypotheses are more conspicuous. These hypotheses respond to mouse-overs by popping up the most likely alternate hypotheses, and those hypotheses replace erroneous ones with a simple click. If an entire action item is rubbish, one click will delete it and provide negative feedback to our models. A user who just wants to make a reasonable action item disappear can click an *ignore this* box, which will still provide positive feedback to our model.

Topics appear as word vectors (ordered lists of words) for direct browsing or to help with user-defined topic queries. Given a user search term, the most likely associated topics are displayed, together with sliders that allow the user to rate the relevance of each list of words to the actually desired topic. As the user rates each topic and its words are re-weighted, a new list of the most relevant words appears, so the user can fine-tune the topic before the browser retrieves the relevant meeting segments.

4 Learning from Feedback

Action Item Feedback. The supervised action item classifiers can be retrained given utterance data annotated as positive or negative instances for each of the utterance classes (task description, timeframe, owner and agreement). User confirmation of a hypothesized action item allows us to take the utterances used to provide its properties as positive instances; conversely, deletion allows us to mark them as negative instances. Switching from one hypothesis to another for an individual property allows us to mark the utterances corresponding to the accepted hypothesis as positive, and the others as negative. Creation of a new action item, or manual editing or insertion of a property value requires us to search for likely utterances to treat as corresponding positive evidence; this can be done by using the relevant sub-classifier to score candidate utterances, and/or by string/synonym comparison, depending on the property concerned. Feedback therefore provides implicit supervision, allowing re-training models for higher accuracy or user-specificity.

Topic Feedback. The topic extraction and segmentation methods are essentially unsupervised and therefore do not need to use feedback to the same degree. Yet even here we can get some benefit: as users define new topics during the search process (by moving sliders to define a new weighted topic mixture), these new topics can be added to the topic pool. They can then be presented to the user (as a likely topic of interest, given their past use) and used in future searches.

References

S. Banerjee, C. Rosé, and A. Rudnicky. 2005. The necessity of a meeting recording and playback system, and the benefit of topic-level annotations to meeting browsing. In *Proceedings of the 10th International Conference on Human-Computer Interaction*.

M. Purver, P. Ehlen, and J. Niekrasz. 2006. Detecting action items in multi-party meetings: Annotation and initial experiments. In *Proceedings of the 3rd Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithms*.

S. Tucker and S. Whittaker. 2005. Accessing multimodal meeting data: Systems, problems and possibilities. In S. Bengio and H. Bourlard (Eds.), *Machine Learning for Multimodal Interaction: First International Workshop, 2004*, v. 3361 of *Lecture Notes in Computer Science*, 1–11. Springer-Verlag.

Scene-Sentence Integration:
Incremental Effects of Mismatch and Scene Complexity

Pia Knoeferle
Dept. of Computational Linguistics
Saarland University, Germany
knoeferle@coli.uni-sb.de

Monica Rodriguez
Dept. of Computational Linguistics
Saarland University, Germany
monic@coli.uni-sb.de

Abstract

We monitored eye movements in a scene during spoken sentence comprehension to investigate the effects of different types of scene-sentence mismatch (action vs. role relations) and of scene complexity on comprehension. Gaze analyses revealed rapid effects of both role relations mismatch and scene complexity, while effects of action mismatch were slightly delayed.

1 Introduction

Verification-task studies have reported longer response latencies (e.g., Just & Carpenter, 1971) and gaze durations (Underwood, Jebbett, & Roberts, 2004) for resolution of a sentence-picture mismatch compared with a match, suggesting a mismatch is more complex to process than a match. We extended the mismatch approach by investigating how different types of scene-sentence mismatch (action versus role relations mismatch, Experiment 1), as well as scene complexity (Experiment 2) affect incremental thematic interpretation. To obtain further insights into the time-course of scene-sentence integration, we monitored participants' eye movements in a scene during comprehension of a related utterance.

2 Experiment 1

2.1 Method

Twenty-four German native speakers with normal vision received each five euro for experiment participation. There were 24 items. Presenting the sentence in Table 1 with the four images in Fig. 1 (A to D) created four conditions (see Table 1).

For counter-balancing reasons, one item had two sentences and four images, resulting in eight experimental lists. Items were rotated across lists such that no participant saw more than one version of each item, and such that each condition appeared equally often in each list. Consecutive experiment trials were separated by at least one of 48 filler trials. An SMI EyeLink I head-mounted tracker monitored participants' gaze in the scene during spoken comprehension. There was no verification task. Rather, participants were instructed to try to understand both sentences and depicted scenes. For half of the 48 filler trials, a written yes/no question about the sentence ensured that people performed a comprehension task. We report analyses of gaze durations that started in the ADV (from adverb onset to the onset of the second noun phrase), and NP2 regions. During these time regions the available scene and utterance information should permit resolution of both the action and role mismatch. If these two types of mismatch rapidly affect thematic interpretation, then their effects should be reflected in the inspection

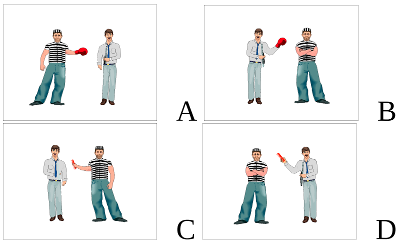


Figure 1: Example Item Images

Sentence & Fig.	Role	Action
1A Der Sträfling boxt gerade den Flötisten 'The convict (S) punches currently the flautist (O)'	Match	Match
1B Der Sträfling boxt gerade den Flötisten 'The convict (S) punches currently the flautist (O)'	Mism.	Match
1C Der Sträfling boxt gerade den Flötisten 'The convict (S) punches currently the flautist (O)'	Match.	Mism.
1D Der Sträfling boxt gerade den Flötisten 'The convict (S) punches currently the flautist (O)'	Mism.	Mism.

Table 1: Example Item Sentences

brandial 2006: Proceedings of the 10th Workshop on the Semantics and Pragmatics of Dialogue, pages 175–176.

durations on the target characters (the scene agent, ‘the convict’, and patient, ‘the flautist’) during the analyses regions.

2.2 Results and Discussion

The key finding is the rapid effect of the role relations mismatch on thematic interpretation as evidenced by an interaction between target character (agent, patient) and role mismatch in the ADV region ($ps < 0.01$, see Fig. 2). People inspected the patient longer than the agent for a role match (C1 & C2, Fig. 2), while there was no such difference for a role mismatch. In contrast, there was no reliable effect of action mismatch in the ADV region. For the NP2 region, there were no reliable effects of the mismatch regarding gaze durations on the target characters.

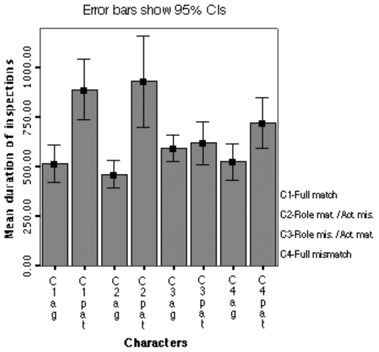


Figure 2: Mean inspection durations to the target characters for the ADV region in Experiment 1

3 Experiment 2

Experiment 2 reused the materials from Experiment 1 but retained only the action mismatch to verify its effects independent of the role relations mismatch. We further examined the influence of scene-complexity (simple vs. complex) on scene-sentence integration. Simple scenes contained the two target characters (agent, patient) of Experiment 1 and four distractor objects. Complex scenes showed an additional three characters.

3.1 Method

Thirty-two further participants from the same population as in Experiment 1 were each paid five euro. Procedure, task, and the analyses regions were the same as in Experiment 1. In addition, we examined early effects of scene complexity by analyzing the duration of inspections that started after NP1 and before verb onset.

3.2 Results and Discussion

There was a main effect of scene complexity for NP1 ($ps < 0.01$), with longer inspection durations on target characters (agent, patient) for simple than complex images. During the ADV region we found no effects of either action mismatch or scene complexity. For NP2, there was an interaction of mismatch and target character ($ps < 0.001$): people fixated the patient longer than the agent for the action-match conditions (C1 & C3). For action-mismatch conditions (C2 & C4), in contrast, inspection duration on the agent and patient did not differ (Fig. 3).

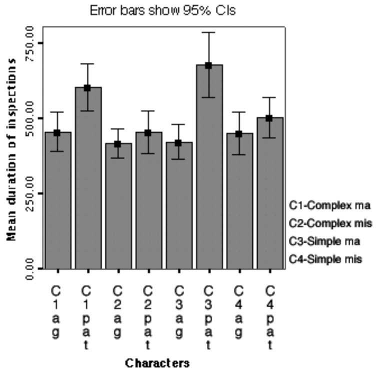


Figure 3: Mean inspection durations to target characters for the NP2 region in Experiment 2

4 Conclusions

Taken together, our findings support the view that scene-sentence integration takes place incrementally. There were, however, differences in the time course of processing actions and role relations mismatch: While the role relations mismatch influenced thematic interpretation post-verbally, effects of the action mismatch only affected thematic interpretation later, during the NP2 region. Scene complexity did not interact with action mismatch, but influenced the inspection duration of the target characters during NP1.

References

Just, M. A., & Carpenter, P. A. (1971). Comprehension of negation with qualification. *Journal of Verbal Learning and Verbal Behavior, 10*, 244–253.

Underwood, G., Jebbett, L., & Roberts, K. (2004). Inspecting pictures for information to verify a sentence: eye movements in general encoding and in focused search. *The Quarterly Journal of Experimental Psychology, 56*, 165–182.

Perspective guides interpretation of questions, declarative questions and statements in unscripted conversation

Sarah Brown-Schmidt*
Department of Brain and Cognitive Sciences
University of Rochester
brownsch@uiuc.edu

Christine Gunlogson
Department of Linguistics
University of Rochester
gunlog@ling.rochester.edu

Duane G. Watson*
Michael K. Tanenhaus
Department of Brain and Cognitive Sciences
University of Rochester

Abstract

This paper describes research investigating the on-line production and interpretation of questions, declarative questions, statements and their replies. Specifically, we examine the role of shared and private knowledge in the processing of these constructions in unscripted conversation. Questions provide a critical test case for the use of perspective in language processing because their felicitous use requires speakers to distinguish common from private knowledge. Analyses of speech and gaze demonstrate that interlocutors distinguish shared from private information and that attention is directed toward different types of entities depending on utterance form. We argue for a central role of perspective in language processing. Discrepancies in experimental findings regarding use of perspective are discussed in terms of relevance of perspective to the task and the utterances of interest.

Cooperative speakers ask questions when they don't know the answer, but believe their addressee might. They assert things they know but believe their addressee might not know. Since Stalnaker's pioneering work on mutual knowledge (Stalnaker, 1978), formal theories of discourse in computational linguistics and within pragmatics and semantics have assumed that keeping track of shared and private commitments and knowledge is central to conversation (Clark, 1992).

While the presuppositions tied to use of different constructions suggest that the distinction between private and shared knowledge is basic to language processing, addressees often fail to distinguish shared from private information (Keysar, Lin and Barr, 2003), and when they do, the egocentric perspective can interfere with reference interpretation (Hanna, Tanenhaus & Trueswell, 2003). However, this and other online work on perspective used imperatives, which may encourage egocentrism due to authority-induced suspension of skepticism and the addressee's aim not to appear confused. Additionally, in order to have control over the interaction and generate specific experimental utterances, these experiments typically employ confederate speakers who are practiced and knowledgeable about the task. However, there is reason to believe that participants interact with confederates differently than they interact with another naïve participant (see Lockridge & Brennan, 2001).

In the experiment described in this paper, we used a goal-directed interactive conversation to examine five semantic-syntactic forms (a-e, see Table 1) that differ in discourse function (requesting/ imparting/ confirming information). Using interactive conversation between naïve participants assures that the constructions are appropriate for the linguistic context and for the knowledge states of the two participants. Thus, speakers will only ask questions when they really don't know the answer, and only make statements when they do. Examining utterance forms which presuppose a distinction between speaker and hearer knowledge (e.g. questions and replies) should provide insights into whether and when this information is used as language is processed on-line.

*Sarah Brown-Schmidt and Duane Watson are now at the Department of Psychology, University of Illinois at Urbana-Champaign.

a	Wh-Question	What's next to the pig with the hat?
b	Statement	There's a cow with shoes next to the pig with the hat.
c	Declarative question	It's a cow with shoes?
d	Question response	(*What's next to the pig with the hat?*)..A cow with shoes.
e	Acknowledgment	(*There's a cow with shoes.*)… A cow with shoes.

We examined the on-line interpretation of wh-questions, declarative questions and statements, and the on-line production of question responses and acknowledgments. Wh-questions and statements were selected to have parallel syntactic structures; each asked about or mentioned the location of one entity (target) with respect to another previously mentioned entity (anchor). If the distinction between shared and private perspectives can be used on-line, we would expect that addressees would direct attention toward private information as they interpret wh-questions, and towards shared or speaker-private information for statements.

Declarative questions, or rising declaratives (Gunlogson, 2001) were used because they have the syntactic form of a declarative, but have question-like intonation and distinct discourse functions. In this task, participants typically used declarative questions to request confirmation or to express skepticism (e.g. *That's a cow with shoes?*). We expected the interpretation pattern for declarative questions to share similarities with both wh-questions and statements.

The question responses and acknowledgments shared a similar syntactic structure (typically a bare noun phrase), however we expected that speakers would direct more attention to private entities when preparing question responses and to shared entities when preparing acknowledgments.

Our results demonstrate that the distinction between shared and private game-pieces is reflected in referent-type differences across utterance forms, and on-line production and interpretation of utterances with different discourse functions.

Wh-questions primarily inquired about addressee-private game-pieces, whereas statements were about shared or speaker-private game-pieces. The pattern of referent-types for declarative questions was half-way between that for wh-questions and statements, with declarative questions primarily inquiring about addressee-private game-pieces and sometimes about shared or speaker-private game-pieces.

When we analyzed the fixations that addressees made as they interpreted these expressions, we saw evidence for a distinct interpretation pattern for wh-questions: Fixations to addressee-private and shared game-pieces were initially equivalent, but following reference to the anchor, addressee-private fixations rose and shared fixations dropped. In contrast, for statements, most fixations were directed to shared game-pieces, suggesting that addressees distinguish shared and private information during on-line interpretation, and direct attention to information relevant for the type of utterance being interpreted.

The relationship between referent type and utterance form confirms our assumptions about the felicity conditions associated with questions and statements. More importantly, using goal-directed conversation and naïve participants, we demonstrated that interlocutors take into account each other's perspective when producing and comprehending utterances for which perspective is relevant. Differences in experimental findings regarding the use of perspective in on-line language processing may be best understood by considering whether perspective was relevant to the task and relevant for interpreting the critical utterances. Continued work using a variety of syntactic structures and communicative situations is needed to understand more precisely when perspective is and is not used in language processing.

References

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

Gunlogson, C. (2001). *True to Form: Rising and Falling Declaratives as Questions in English.* Unpublished Doctoral Dissertation, University of University of California, Santa Cruz, Santa Cruz, CA.

Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, 49, 43-61.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, 89, 25-41.

Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin and Review*, 9, 550-557.

Stalnaker, R. C. (1978). Assertion. In P. Cole (Ed.), Syntax and semantics: *Pragmatics* (Vol. 9, pp. 315-332). New York, NY: Academic Press.