Subdialogs and Attentional State

Donna K. Byron*
Dept of Computer and Info. Science
The Ohio State University
byron.18@osu.edu

Sarah Brown-Schmidt
Dept of Brain and Cognitive Sciences
University of Rochester
sschmidt@bcs.rochester.edu

1 Introduction

Assigning meanings to Noun Phrases (NPs) is crucial to the language understanding process. Previous research demonstrated that many sources of information are used in this process, e.g. task pragmatics (Brown-Schmidt et al., 2002), expectations introduced by disfluencies (Arnold et al., In press), and the task-relevant properties of objects (Chambers et al., 2002). (Grosz and Sidner, 1986) (henceforth G&S) showed that the relative embedding of discourse segments also affects the meaning of NPs. This report presents the initial findings of a psycholinguistic experiment to examine whether subjects are able to utilize this high-level property, discourse structure, in the earliest moments of NP interpretation.

The model reported in G&S organizes the entities in a discourse as a set of *focus spaces*, one per discourse segment, that are managed as a stack. This stack, called the *attentional state*, are searched to find referents for anaphoric NPs. As a discourse segment unfolds, its focus space accumulates discourse referents. When the discourse segment ends, its entities are removed from the stack. In the G&S model, 'only the attentional state can constrain the interpretation of referring expressions directly'[G&S, pg. 180].

G&S is not intended as a cognitive model; however, the theory leads to an intriguing question for human processing. After the close of a subdialog that interrupts a discourse segment, are entities from that subdialog considered, if only briefly, as the referents for subsequent NPs or are they 'popped off the stack' and not considered?

2 Experimental Design

To examine this question, we conducted a preliminary experiment using the visual-world paradigm, in which we obtained a record of the participants' eye-fixations with the use of a head-mounted eye-tracker. Approximately 200ms after the onset of an NP, a participant will look at items he is considering as the referent (Tanenhaus et al., 1995). The results we describe here must be considered to be preliminary since the experiment included only one trial. Data was collected from 36 undergraduate participants[1] from the University of Rochester.

Participants sat at a table with supplies for dying eggs, including a carton containing either 1 egg & 1 non-egg competitor object or 2 eggs. The participants then heard pre-recorded instructions to dye an egg, shown in Table 1. The instructions for each condition include an initial instruction (segment A), a subdialog (segment B) directing subjects to prepare a bowl of dye and arrange objects on the table, and a final segment after the subdialog closes that returns to the main task of dying the egg (segment C).

Subjects were randomly assigned into one of four conditions. Conditions varied along the following two dimensions: 1) the competitor object is an egg or a non-egg (a small sheet of stickers used to decorate eggs), 2) the competitor object is or is not mentioned in the subdialog. The non-egg competitor condition serves as a baseline for comparing the proportion of looks to the competitor egg (conditions C1 and C2) with the number of looks to the stickers (conditions C3 and C4).

The critical point in the experiment is the

*This material is based upon work supported by the National Institutes of Health under award number NIH HD-27206 to M.K. Tanenhaus.

[1] 10 in conditions C1 and C3; 9 in conditions C2 and C4. Audio for 1 person in each of C2 and C4 was lost.

Table 1
Instruction Sequences

C1) Competitor egg is mentioned in subdialog	**C3) Competitor stickers are in subdialog**
A1) In this experiment you're going to be dying an easter egg.	A1) In this experiment you're going to be dying an easter egg.
A2) Pick which egg you want to paint.	A2) *Put the egg in the egg stand so that you can paint it.*
A3) *And put it in the egg stand so you can paint it.*	
B4) Now before we can proceed we need to prepare the dye.	B3) Now before we can proceed we need to prepare the dye.
B5) Fill up the empty bowl with water.	B4) Fill up the empty bowl with water.
B6) Pick which color of dye you want to use.	B5) Pick which color of dye you want to use.
B7) Add a couple of drops of dye to the water.	B6) Add a couple of drops of dye to the water.
B8) *Move the styrofoam carton with* **the egg** *in it over by the towel and out of the way.*	*B7) Move the styrofoam carton with the stickers in it over by the towel and out of the way.*
C9) Great! Now that's all done so we're ready to dye.	C8) Great! Now that's all done so we're ready to dye.
C10) Use the dipper to pick up **the egg** and dip it in the dye.	C9) Use the dipper to pick up **the egg** and dip it in the dye.
C2) Competitor egg is not mentioned in subdialog	**C4) Competitor stickers are not in subdialog**
A1) In this experiment you're going to be dying an easter egg.	A1) In this experiment you're going to be dying an easter egg.
B2) First we need to prepare the dye.	B2) First we need to prepare the dye.
B3) Fill up the empty bowl with water.	B3) Fill up the empty bowl with water.
B4) Pick which color of dye you want to use.	B4) Pick which color of dye you want to use.
B5) Add a couple of drops of dye to the water.	B5) Add a couple of drops of dye to the water.
B6) *Pick which egg you want to paint and put it in the egg stand so you can paint it.*	B6) *Put the egg in the egg stand so that you can paint it.*
C7) Great! Now that's all done so we're ready to dye.	C7) Great! Now that's all done so we're ready to dye.
C8) Use the dipper to pick up **the egg** and dip it in the dye.	C8) Use the dipper to pick up **the egg** and dip it in the dye.

mention of 'the egg' in the last instruction. At the onset of this instruction, both the task and the discourse structure indicate a return of attention to the target egg. In C1 and C2, there are two eggs that could be the referent of the critical NP. In C1, the competitor egg (e.g. the egg the participant is not dyeing) was mentioned most recently in the subdialog, but in C2, it was never mentioned. If entities in the subdialog have been 'popped off the stack' in a manner analogous to the G&S model, subjects in C1 should be no more likely to look at the competitor egg than the stickers in C3. If participants consider the competitor egg as a possible referent, we expect to see a larger proportion of fixations to the competitor egg than to the stickers.

3 Results

After hearing the critical instruction, each participant quickly selected the target egg and completed the task. This is consistent with the predictions of G&S, that closing the subdialog eliminated the competitor egg as a plausible referent of the critical NP.

Eye-fixations suggest that closing the subdialog largely (but not entirely) eliminates referents introduced in the subdialog as possible referents of upcoming NPs. After the critical NP "the egg", when the competitor object was the stickers (C3 and C4), participants looked primarily at the target egg and the instrument they are about to use (the dipper), with minimal looks to the stickers. Likewise when the competitor object was an egg, but it was not mentioned (C2), eye fixations followed this same pattern.

Crucially, in C1 we did observe a few looks to the competitor egg. The lack of fixations to the competitor egg in C2 suggests that these fixations were a result of the competitor egg being mentioned in the subdialog. This suggests that entities mentioned in subdialogs may be maintained as potential referents for upcoming NPs.

References

J.E. Arnold, M. Fagnano, and M.K. Tanenhaus. (In press). Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research*.

S. Brown-Schmidt, E. Campana, and M. Tanenhaus. 2002. Reference resolution in the wild. In *24th Annual Meeting of the Cognitive Science Soc.*, Fairfax, VA, August.

C. Chambers, M. Tanenhaus, K. Eberhard, G. Carlson, and H. Filip. 2002. Circumscribing referential domains in real time language comprehension. *J. Memory and Language*, 47:30–49.

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

M.K. Tanenhaus, M.J. Spivey-Knowlton, K.M. Eberhard, and J.E. Sedivy. l995. Integration of visual and linguistic information in spoken language comprehension. *Science*, 268:1632–1634.