The influence of cognitive load on repeated references in speech and gesture

Ingrid Masson-Carro (i.masson@uvt.nl)

Martijn Goudbeek (m.b.goudbeek@uvt.nl)

Emiel Krahmer (e.j.krahmer@uvt.nl)

Tilburg centre for Cognition and Communication (TiCC), School of Humanities,
Tilburg University, The Netherlands.

Abstract

Shared common ground between interaction partners has been found to lead to reduction in repeated references to a target entity, both in speech and gesture. It has been shown, however, that increasing the cognitive load of speakers has the potential to affect how speakers and addressees adapt to one another in dialogue. This paper reports on an experiment in which native speakers of Dutch engaged in a director-matcher task where repeated references were elicited, and a time constraint was imposed in order to increase the load of speakers. Our results show that cognitive load was not an impediment to the reduction process, although it did have an effect on the overall task performance, suggesting that reduction results from rather automatic processes.

Keywords: cognitive load, reduction, referring expressions, speech, gesture.

Introduction

Repeated references in speech and gesture

When people speak, they constantly refer to objects in the environment, often by producing referring expressions such as "the large blue block". It is natural that during the course of a conversation speakers need to refer more than once to a target entity. These consecutive references are known as "repeated references". Repeated references have been found to be reduced in a number of aspects with respect to initial references, both in speech and in speech-accompanying gestures. One likely candidate to influence this reduction is the "common ground" shared by interaction partners: by relying on mutual knowledge, speakers tend to simplify their contributions to a conversation, thus making communication more efficient (Brennan & Clark, 1996).

With regard to speech, studies employing referential paradigms have shown that when some common ground is shared between interlocutors, people use fewer words to communicate (Clark & Wilkes-Gibbs, 1986) and the amount of information contained in the verbal utterances decreases (Fussell & Krauss, 1992). Likewise, words in repeated references are shorter and less articulatory precise -thus less intelligible to naïve listeners than words in initial references (Bard et al., 2000). When we examine the gestures that accompany speech, a similar picture emerges. When information is shared between interaction partners, speakers produce fewer gestures, and these are perceived as less precise and informative than when there is no common ground (Gerwing & Bavelas, 2004; Holler & Wilkin, 2009). Hoetjes et al. (2012) carried out a first series of experiments specifically designed to explore reduction in gesture during the production of repeated references. Their study showed that speech and gesture reduced together, and gestures in repeated references were judged as less precise and more difficult to interpret than gestures accompanying first descriptions of an item. Remarkably, when no addressee was visible, there was no reduction, confirming common ground's decisive role on the reduction process.

Reduction and cognitive load

"Reduction" seems to be ubiquitous in repeated mentions to a same item, thus one may argue that it is an inherent feature of dialogue. In this paper we explore one of the factors potentially mediating reduction, namely cognitive load. In dialogue, speakers rely on the information they share with their addressees when building their utterances, but this feature seems to be offset when speakers are under cognitive load (Horton & Keysar, 1996; Goudbeek & Krahmer, 2011). This suggests that increasing cognitive load can present a barrier to audience design. However, it has not been tested how cognitive load might affect reduction. The only study exploring reduction and cognitive load that we are aware of is that by Howarth and Anderson (2007), in which speakers and addressees had to participate in a referential collaborative task whilst being subject to a time-pressure constraint. In their study, articulatory reduction in repeated mentions took place irrespective of cognitive demands. Taking these seemingly contradictory results altogether, it seems plausible that two different mechanisms are at play, in line with Bard et al's (2000) Dual-Process model, which posits that in dialogue some aspects are cheap and occur automatically (e.g. articulatory reduction) whereas other aspects require conscious involvement and are thus costly (e.g. making explicit use of common ground knowledge). Nonetheless, Howarth and Anderson only explored reduction at the word level, so it remains to be seen whether reduction occurred at other levels, such as the syntactic level (e.g. number of words). Thus, it might also be the case that other types of reduction occur regardless of load, which would be in line with models that consider alignment between interaction partners as an automatic process, the result of a priming mechanism operating at different levels (Pickering & Garrod, 2004).

Until now, no studies have paid attention to how cognitive load affects gesture reduction, and in general research on the relationship between gestures and cognitive load has yielded mixed results. On the one hand, gesture has been argued to reduce cognitive load for the speaker, e.g. by facilitating speech planning (Kita, 2000). In accordance with this, studies that manipulated the amount of load have shown that speakers produce more gestures during verbal difficulties, while describing a complex referent, or when faced with a dual task (Morsella & Krauss, 2004; Melinger & Kita, 2007). On the other hand, we have studies such as de Ruiter et al. (2012), who found that the rate of iconic gestures is not affected by whether images are easy or hard to describe, or Mol et al. (2009), who found that an addressee-tailored use of co-speech gestures may come at a higher cognitive cost to the speaker, instead of the opposite.

In summary, our knowledge of how cognitive load affects the processes underlying dialogue is limited and the present study investigates how cognitive load affects the reduction of multimodal repeated references. If our results show that cognitive load does not affect reduction, we might conclude that reduction is a rather automatic process that does not require cognitive resources to operate. If, on the other hand, reduction were influenced by cognitive load, it could imply that we are facing a rather costly process. Both these outcomes will provide important information for already-existing dialogue models such as Pickering and Garrod's (2004) and Bard et al.'s (2000).

Method

In this experiment, participants were recorded during the completion of a director-matcher referential task in which repeated references to a series of target objects were elicited. The experiment followed a mixed design, with repetition as the within-subjects variable and cognitive load (time pressure) as the between-subjects variable.

Participants

Eighty-two students from Tilburg University (*M* = 21.1; *SD* = 5.85) took part in this experiment, in exchange for course credit points. Participants carried out the experimental task in pairs, therefore data from forty-one dyads were collected. All participants were native speakers of Dutch.

Stimuli

The stimuli materials consisted of four monochrome sets of abstract pieces: a green, a red, a blue, and a yellow one. Each set was built using a range of Lego and Duplo blocks of varying sizes and shapes, plus a series of abstract composite (indivisible) figures specifically built for the task. Of these composite figures, we selected two target pieces per set, summing a total of eight target pieces (Figure 1).

Twelve models were created (three per colour set). Each model was composed of the two target pieces of the corresponding set plus two filler objects. To guide the directors through the task, a program was written using ActionScript 5.0, where step by step instructions were provided, both for the referential (target piece retrieval) and the instruction (model assembly) tasks.

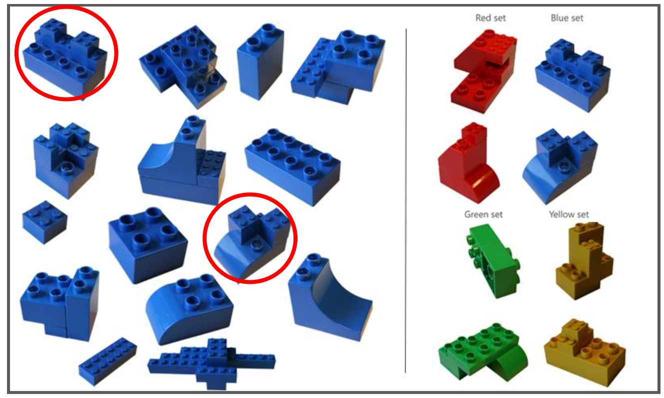

Figure 1: (right) the 8 target; (left) example of all pieces contained in one of the four Lego sets (blue). The two target pieces are circled in red.

Procedure

The experiment was conducted in a lab. Upon their arrival to the experiment room, participants were randomly assigned the roles of director and matcher and sat across a table. The setup (Figure 2) was arranged in such a way that both participants would have visual access to the working space, but the directors could not see the buckets by the matcher's side, and the matchers could not see the director's screen. Both participants received written instructions, and could ask questions if anything was unclear. Participants were encouraged to interact freely during the experiment and collaborate as much as required in order to assemble the twelve models correctly.

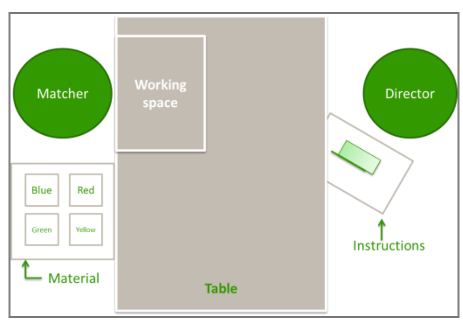

Figure 2: Experimental setup

Each dyad had to accomplish twelve trials (three trials per Lego set), the order of which was semi-randomised. The first part of a trial consisted of a referential task, where the director was asked to describe four pieces to the matcher (the two target pieces, plus two fillers), who had to retrieve those pieces from one of the buckets and position them on the working space. Once all the pieces were successfully retrieved, the director would press a button on the computer to proceed with the second part of the trial, where the director had to instruct to matcher on how to assemble a model with the pieces retrieved.

Participants in the "no cognitive load" condition could devote as much time as needed to the task, whereas participants in the “cognitive load” condition had 120 seconds to accomplish each trial (for the retrieval and assembly tasks combined). The length of this period was established during pilot research and was implemented by means of a timer present on the screen of the instructor, counting down from 120 to 0. When 0 was reached, participants would be directed to the next trial automatically. Therefore, the objective was to retrieve the pieces as quickly as possible in order to have time left to assemble the model.

Data analyses

Speech Speech was transcribed verbatim, and all first and third mentions to the target pieces were selected. These references were analysed in terms of their overall duration in msec, number of words, and level of information. The latter factor was accounted for by the occurrence of meaningful units in the speech, coded as “attributes”. Based on all the director’s descriptions we configured a list of eight attributes that were consistently used to describe the blocks: size (e.g. “small”), shape (e.g. “oval”), position (“above”), simple element (e.g. “block”), specific element (e.g. “building”, “city-wall”), colour, number, and other informative adjectives (e.g. “complex”).

Gesture All gestures accompanying first and third mentions were identified. First, the number of gestures, the gesture rate (no. gestures / no. words) and the duration of the gestures were determined. Second, gesture size was annotated on a five-point scale that judged the size of the stroke from small (1) to big (5). Additionally, we annotated whether a gesture was performed with one, or the two hands.

Statistical analyses The main statistical procedure was a Repeated Measures ANOVA, with “repetition” (two levels) and “target piece” (eight levels) as the within-subjects variables, and “cognitive load” as the between-subjects variable. An additional ANOVA was run to evaluate the performance of the participants on the referential task.

Results

The referential task generated a total of 307 initial references, and 300 repeated (third-mention) references. The results confirmed an effect of cognitive load on the participant’s performance (i.e. the mean time needed to retrieve all four pieces per trial *correctly*) [F (1, 20) = 35.56, p <.001], with performance times in initial references lasting longer for participants without load (*M* = 90563, *SD* = 12616) than for participants under load (*M* = 54521, *SD* = 11931). In repeated references, no such effect was found, suggesting that participants with no time constraints became more efficient at the task than participants with limited time (see Figure 3).

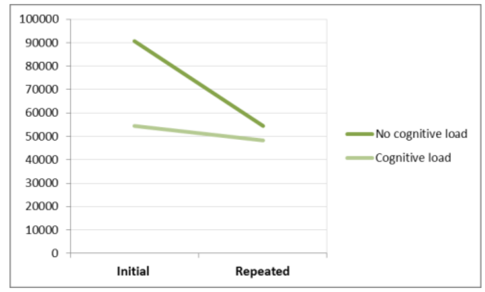

Figure 3: Overall performance of participants in both conditions, in initial and repeated references.

Our results indicate that there was no main effect of cognitive load on the reduction process, for speech nor for speech-accompanying gestures. However, we do find an interaction between repetition and cognitive load for speech duration [F (1, 39) = 5.95, p <.05] and mean number of words per reference [F (1, 39) = 4.5, p <.05] (Figure 4).

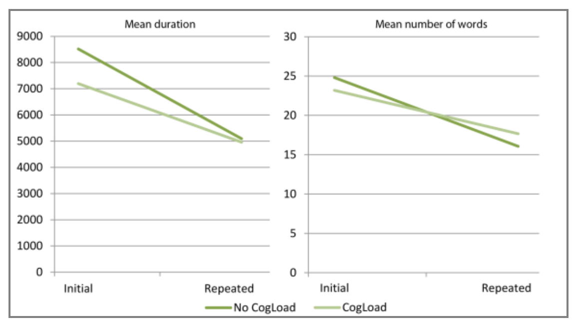

Figure 4: Interaction effects between “repetition” (reduction) and “cognitive load”

Reduction in speech

In speech, reduction in repeated references was significant for all the variables analysed (Table 1), that is, the mean duration of the referring expressions, mean number of words, and the mean number of semantic attributes per reference [F (1, 39) = 162.7, p <.001].

Table 1: Means and Standard Deviations of initial and repeated references in speech.

	Initial (*SD*)	Repeated (*SD*)
Duration in msec*	7939 (*2265*)	5034 (*1506*)
No. words*	24.1 (*7.02*)	16.76 (*4.85*)
No. attributes*	5.92 (*1.56*)	4.26 (*1.18*)

* *p <.001*

Reduction in gesture

An effect of repetition was found for the number of gestures [F (1, 39) = 23.73, p <.001], indicating that speakers produced less gestures in third mentions to an object (*M* = 2.44, *SD* = 0.81) than in their initial descriptions (*M* = 2.98, *SD* = 1.24). Nevertheless, repetition had no significant effect on the other variables analyzed (gesture duration, gesture size, and number of two-handed gestures).

Discussion

The results reported in this paper have shown that cognitive load, elicited by means of time pressure, does not hinder the reduction of repeated references, in speech nor in gesture – although it does affect overall task performance. For speech, both syntactic and semantic reduction still took place under cognitive load. For gesture, only the number of gestures was reduced with repetition, replicating the results of Hoetjes et al. (2012). Our results did show an interaction between cognitive load and reduction, due to participants under cognitive load producing shorter first references. Interestingly enough, none of the variables analyzed show differences across conditions with respect to third references, indicating that both groups reached the same end-level of reduction. Thus, we conclude that, whereas time pressure affected the initial mentions to a target entity, the reduction process remained unaffected.

Our finding that reduction takes place irrespective of cognitive load is in line with the findings by Howarth & Anderson (2007) and complies with the view exposed in Pickering and Garrod's Interactive Alignment Model (2004), suggesting that reduction may be the result of a priming mechanism that is resource-free and operates automatically at different levels. Nevertheless, we would like to argue for another plausible explanation. It is likely that different types of cognitive load might elicit different communicative patterns. This is suggested by the fact that all the studies where cognitive load has an effect on dialogue have employed dual-task paradigms (intrinsic load), with the exception of Horton and Keysar (1996). Remarkably, studies employing time pressure to elicit load (extrinsic load), showed that reduction occurred despite the manipulation (e.g. Howarth & Anderson, 2007).

Likewise, most studies dealing with gesture and cognitive load have manipulated task difficulty or have given their participants two simultaneous tasks to accomplish (e.g. Melinger & Kita, 2007), rather than subjecting participants to time pressure. We argue that it is important to contrast the communicative patterns generated by cognitive load as elicited by different paradigms in order to get a more complete picture of the mechanisms at play in dialogue.

Acknowledgments

We thank Kristel Bartels for the help in transcribing the Dutch dialogues. We received support through the *Koninklijke Nederlandse Akademie van Wetenschappen* (KNAW), via a grant awarded to the first author.

References

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, 42, 1-22.

Brennan, S., & Clark, H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology*, 22, 1482-1493.

Clark, H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, 22, 1-39.

de Ruiter, J. P., Bangerter, A., & Dings, P. (2012). The interplay between gesture and speech in the production of referring expressions: Investigating the trade-off hypothesis. *Topics in Cognitive Science*, 4, 232–248.

Fussell, S. R., & Krauss, R. M. (1992). Coordination of knowledge in communication: Effects of speakers' assumptions about what others know. *Journal of Personality and Social Psychology*, 62, 378-391.

Gerwing, J., & Bavelas, J. (2004). Linguistic influences on gesture's form. *Gesture*, 4, 157-195.

Goudbeek, M. & E. Krahmer (2011). Referring under load: Disentangling preference-based and alignment-based content selection processes in referring expression generation. *Proceedings of the Workshop on Production of Referring Expressions* (Pre-CogSci 2011). 20 July 2011, Boston, Massachusetts.

Hoetjes, M., Koolen, R., Goudbeek, M., Krahmer, E. & Swerts, M. (2012). Reduction in speech and gesture during the production of repeated references. Manuscript submitted for publication.

Holler, J., & Wilkin, K. (2009). Communicating common ground: how mutually shared knowledge influences speech and gesture in a narrative task. *Language and Cognitive Processes*, 24, 267-289.

Horton, W. & Keysar, B. (1996). When do speakers take into account common ground? *Cognition* 59, 91-117.

Howarth, B. & Anderson, A.H. (2007). Introducing objects in spoken dialogue: the influence of conversational setting and cognitive load on the articulation and use of referring expressions. *Language and Cognitive Processes*, 22, 272-296.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge, England: Cambridge University Press.

Melinger, A., & Kita, S. (2007). Conceptualisation load triggers gesture production. *Language and Cognitive Processes*, 22, 473–500.

Mol, L., Krahmer, E., Maes, A. & Swerts, M. (2009). Communicative gestures and memory load. In N.A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 1569-1574). Austin, TX: Cognitive Science Society.

Morsella, E., & Krauss, R. (2004). The role of gestures in spatial working memory and speech. *The American Journal of Psychology*, 117, 411–424.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169-225.