On the automaticity of reduction in dialogue: Cognitive load and repeated multimodal references

Ingrid Masson-Carro (i.massoncarro@tilburguniversity.edu)
Martijn Goudbeek (m.b.goudbeek@tilburguniversity.edu)
Emiel Krahmer (e.j.krahmer@tilburguniversity.edu)

Tilburg center for Cognition and Communication (TiCC), Tilburg University, The Netherlands

Abstract

In conversation, speakers are likely to refer to the same objects more than once. These repeated references are reduced with respect to their initial counterparts, both in speech and gestures. In this paper we investigate the effect of cognitive load on the reduction of multimodal referring expressions. We report 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. Our results show that articulatory, lexical, semantic, and gestural reduction took place irrespective of the cognitive demands. Nevertheless, we found that cognitive load moderated the extent to which these utterances were reduced, with reduction being less pronounced for speakers experiencing higher load. A subsequent perception experiment revealed that speakers with an increased load produced referring expressions that proved more informative to naïve listeners.

Keywords: Reduction, referring expressions, gesture, cognitive load

Introduction

In face-to-face dialogue, speakers often produce referring expressions (e.g., "that large red block") to talk about objects that are present in their immediate, shared physical context. These expressions do not only consist of speech, but may also include hand gestures that complement or emphasize what is being said in words (e.g. saying "that large red block" -while tracing a shape in the air). When speakers engage in conversation, it is likely that they mention some objects more than once. These subsequent references are known as "repeated references" or "repeated mentions". Previous research has shown that referring several times to an object results in reduction of the repeated mentions, both regarding speech (e.g., Clark & Wilkes-Gibbs, 1986; Bard & Aylett, 2004) and co-speech gestures (e.g., Hoetjes et al., 2011). Nevertheless, the dialogical context and the role of the addressee strongly mediate the extent to which speakers reduce their utterances. For example, reduction might be suppressed when retelling the same story to a new (naïve) listener with whom no *common ground* is shared (Galati & Brennan, 2010), or enhanced after receiving positive feedback from the addressee (Holler & Wilkin, 2011). In this study we look at the influence of cognitive load on the reduction of repeated referring expressions, to find out more about how speakers and addressees communicate in moments of high load. Concretely, we ask whether reduction is facilitated by automatic processes that mainly confer cognitive benefits to the speakers, or whether reduction stems from more cognitively demanding processes, helping to make utterance processing easier for the listeners.

When do we reduce *what*, and *why*?

Reduction in repeated references has been consistently observed at the acoustic, lexical, syntactic, semantic, and discourse levels, and also in the number and appearance of the gestures that accompany speech. In speech, words in repeated references are shorter and less articulatory precise than words in initial references (Bard et al., 2000; Bard & Aylett, 2004); there is a decrease in the number of words contained in the reference phrases (Krauss & Weinheimer, 1964; Clark & Wilkes-Gibbs, 1986), and also in their semantic content (Hoetjes et al., 2011). Furthermore, repeated references are accompanied by fewer co-speech gestures than initial references, and these gestures are less precise, and in some cases smaller, than their first-mention counterparts (Gerwing & Bavelas, 2004; Holler et al., 2011; Hoetjes et al., 2011).

This reduction seems rather natural: it would be hard to maintain a conversation in which each object is referred to every single time with a full description. In this light, reduction can be seen as an instrument that contributes to the fluidity of our communicative exchanges with others. However, the mechanisms underlying reduction are less clear. For example, what is the degree of audience design involved in reduction? Do speakers reduce their utterances because it is easy for themselves (speaker-oriented benefits), or for the sake of more successful communication with their interlocutors (addressee-oriented benefits)?

One hypothesis is that reduction is tightly coupled with the knowledge that accrues between interlocutors as the conversation unfolds, known as *common ground* (e.g., Clark & Wilkes-Gibbs, 1986; Brennan & Clark, 1996; Galati & Brennan, 2010). Assuming that the goal of referential communication is for an addressee to identify a target, it is plausible that when a target is often referred to, fewer words are needed to achieve understanding. Compatible views argue that already expressed information becomes "more predictable" in conversation and is therefore shortened (e.g., Levy & Jaeger, 2007). Early studies show that acoustic reduction (i.e., faster articulation of words) does not occur, or occurs at a slower pace, when there is no addressee (Krauss & Weinheimer, 1964), or in a decontextualized setting –e.g., repeating lists of words (Fowler, 1988). Similarly, retelling a story to a same (old) addressee results in acoustic, lexical and semantic reduction of the utterances, whereas retelling the same story to a new addressee does not (Galati & Brennan, 2010). Thus, it is safe to say that the communicative setting plays an important role in mediating reduction. However, repeated references might also become reduced simply because their antecedent is more "accessible" in the speaker's memory (e.g., Ariel, 1990), making retrieval easier, which is in turn associated with faster articulation (Lam & Watson, 2010). Some studies have supported this view. For example, Bard et al. (2000) found that words in repeated mentions were shorter and less intelligible than words in initial mentions, regardless of whether they had been produced towards a new or an old addressee. This opens the discussion on the extent to which reduction is mediated more strongly by speaker-internal or speaker-external (contextual) constraints.

Previous research suggests the existence of two types of processes at play in dialogue, namely fast automatic priming processes that mainly confer benefits to the speaker, and slower processes that might be more cognitively costly – such as dual process model was originally proposed by Dell and Brown (1991), and later by Bard et al. (2000). One way to tap into these dialogue processes is by manipulating the degree of cognitive load experienced by speakers, based on the premise that when the load experienced is high, processes that take more cognitive resources to operate will suffer. Several studies employing cognitive load paradigms have shown that audience design (i.e., adapting to one's addressee during conversation, for example by making use of shared knowledge) seems to be offset when speakers are under high cognitive load (e.g., Horton & Keysar, 1996), even in cases where taking the perspective of the listener would be appropriate, for example when instructing a child how to perform a task (as opposed to an adult) (Roßnagel, 2000). In other words, increasing cognitive load can present a barrier to audience design. Given that reduction in referential communication largely depends on the quality of the interaction with the addressee, it is possible that cognitive load may affect the reduction process. The only study exploring reduction and cognitive load that we are aware of is that by Howarth and Anderson (2007), who asked speakers and addressees 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, suggesting that it is an automatic process, related to, but separate from, addressee adaptation. It remains to be seen whether this result holds for aspects of speech production beyond articulation and, importantly, whether and how cognitive load affect the use of gestures in repeated mentions. So far, most studies dealing with cognitive load only looked at the gesture rate, which is the proportional use of gestures with respect to speech, yielding 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), but can also increase it, if these gestures are communicatively intended (Mol et al., 2009).

The present study

Our knowledge of how cognitive load affects the processes underlying dialogue is limited. Previous studies suggest that reduction is heavily mediated by the interaction with an addressee (e.g., Krauss & Weinheimer, 1964), but we also know that increasing the cognitive load in speakers can present a barrier to audience design (e.g., Horton & Keysar, 1996; Goudbeek & Krahmer, 2011). This leads to the hypothesis that if audience design is affected by increasing cognitive load, reduction (as a form of audience design) might as well be, unless reduction stems from more automatic processes designed to confer cognitive benefits to the speaker. In the present study, our main goal is to investigate whether (and how) cognitive load affects the reduction of multimodal repeated references. Most studies (e.g., Howarth & Anderson, 2007; Bard et al. 2000) have looked at articulatory reduction only (word-length, intelligibility), but repeated references to objects are also lexically, semantically and gesturally reduced. Thus, a comprehensive analysis needs to widen the scope and include all the levels at which reduction has been found to occur in speech and gestures. Our study attempts at performing such an analysis.

Experiment I: production

Method

In Experiment I, participants completed a director-matcher task in which repeated references to a series of eight target objects were elicited. The experiment followed a mixed design, with repetition as the within-subjects variable, and cognitive load (operationalized as time pressure –see Howarth & Anderson, 2007) as the between variable.

Participants Eighty-two students from Tilburg University (M = 21.1 years; SD = 5.85, 46 female, 36 male), all of them native speakers of Dutch, took part in this experiment, in exchange for course credit. Participants carried out the experimental task in pairs, therefore data from forty-one dyads were collected.

Stimuli The materials consisted of four monochrome sets of abstract pieces: a green, a red, a blue, and a yellow set. Each consisted of single Lego and Duplo blocks, and of composite pieces built specifically for the task by gluing together various single pieces to form complex shapes. Of these composite figures, we selected two target pieces per color set, summing to a total of eight pieces that the speakers would have to describe to the matchers (Figure 1). Additionally, using these pieces, we created twelve models (three per color set) that the matchers would ultimately assemble (see: procedure). The directors were guided through both tasks by computerized written step-by-step instructions.

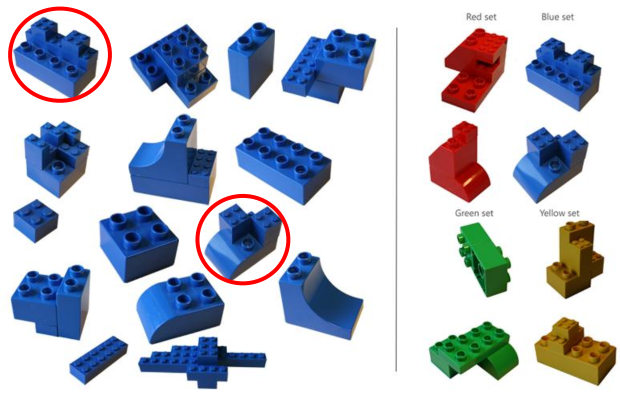

Figure 1: (right) the 8 targets; (left) example of all pieces contained in one of the four Lego sets (blue). The two target pieces are circled in red.

Procedure Participants were randomly assigned the roles of director and matcher, and sat at opposite sides of a table (Fig. 2). Both participants had visual access to the working space, but the matchers could not see the director's screen.

Each dyad had to accomplish twelve semi-randomized trials. Each trial corresponded to the assembly of one of the twelve models, and consisted of two parts. The first part of a trial was the target piece retrieval task, where the director was asked to describe four pieces (the two target pieces, plus two fillers) to the matcher, who had to retrieve those pieces from one of the buckets by her side and position them on the working space. Thus, this manipulation elicited twenty-four key references per speaker (three references per target piece). 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 the matcher on how to assemble a model with the pieces retrieved.

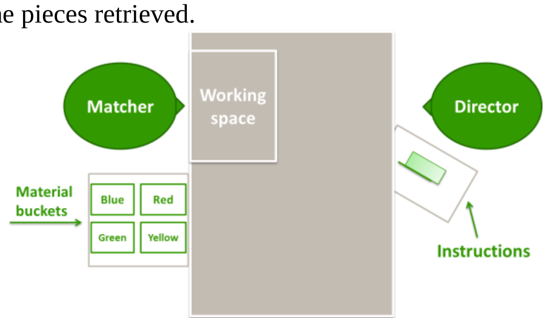

Figure 2: Experimental setup

Participants in the "low load" condition could devote as much time as needed to the task, whereas participants in the "high load" condition had 120 seconds to accomplish each trial (for both tasks). 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 were 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 Verbatim transcriptions of the first, second, and third mentions to the target pieces were selected from the retrieval task. These references were annotated in terms of their duration in msec, number of words, and word duration (in msec). We also performed analyses to explore the type of information contained in the referring expressions. We looked at two variables: semantic content and common ground. To measure the semantic content, we annotated the occurrence of meaningful units in the speech, coded as "attributes". Based on all the director's descriptions we configured a list of attributes that were consistently used to describe the blocks, such as size (e.g., "small"), shape (e.g., "oval"), position ("above"), etc. To measure common ground, we created a scheme to evaluate whether speakers' descriptions took into account the addressee's knowledge and perspective. We considered a referring expression as making use of some basic common ground information when the speaker would mention a piece as an already known one (e.g., "remember the piece you just retrieved? Take it again"), when the speaker would refer to elements in the working space available to both (e.g., "take the piece in front of you, left side of X"), or when a conceptual pact was created (Brennan & Clark, 1996) (e.g., "take *the castle*").

Gesture All iconic gestures (McNeill, 1992) accompanying the referring expressions were identified. First, the number and the duration of the gestures were determined, and we computed the gesture rate (number of gestures in proportion to words). Then, gesture size was annotated on a five-point scale that judged the size of the stroke from small (1) to big (5). We also annotated whether a gesture was performed with one or two hands, and whether there was repetition of the gestural stroke (e.g., tracing the same shape repeatedly).

Statistical analyses The statistical procedure was Repeated Measures ANOVA, with "repetition" (three levels) and "target piece" (eight levels) as the within-subjects variables, and "cognitive load" as the between-subjects variable.

Results

The referential task generated a total of 884 referring expressions. Our analyses show that speakers produced shorter referential phrases when referring to an object for the second and third time, than for the first time [F (2, 78) = 73.15, p <.001, η_ρ^2 = .65]. Likewise, repeated mentions contained fewer words [F (2, 78) = 59.03, p <.001, η_ρ^2 = .6], and these words were articulated faster (i.e., had a shorter duration) than words contained in initial references [F (1, 82) = 9.51, p <.005, η_ρ^2 = .1]. Complementing these results, our semantic analysis reveals that repeated references contained a lower amount of semantic attributes than initial ones [F (2, 78) = 37.37, p <.001, η_ρ^2 = .49]. Lastly, repetition also led to an increase in the use of common ground information [F (2, 78) = 45.2, p <.001, η_ρ^2 = .53], which is consistent with previous research (e.g., Clark & Wilkes-Gibbs, 1986) (see Table 1).

Table 1: Mean values of the dependent speech and gesture variables (for first, second and third references).

Speech	First (*SE*)	Second (*SE*)	Third (*SE*)
Duration (msec)*^+	7901 *(328)*	6020 *(235)*	5074 *(214)*
Articulatory (msec)*^	462.88 *(10.7)*	x	419.85 *(9.88)*
No. words*^+	24.01 *(.99)*	18.92 *(.64)*	16.75 *(.75)*
No. Attributes*^+	5.91 *(.21)*	4.78 *(.2)*	4.29 *(.17)*
Common Ground(%)*^	20.98 *(2.51)*	50.4 *(3.3)*	51.44 *(3.12)*
Gesture			
References with gesture (%)*	78.64 *(3.41)*	62.69 *(4.32)*	69.42 *(4.3)*
No. Of gestures*^	2.6 *(.09)*	2.28 *(.05)*	2.38 *(.07)*
Duration (msec)*+	1094 *(33)*	1533 *(48)*	994 *(23)*
Gesture rate (%)+	8.04 *(.64)*	7.05 *(.65)*	8.89 *(.81)*
Size (Scale 1 - 5)*+	2.69 *(.05)*	2.86 *(.05)*	2.6 *(.04)*
Two-handed (%)	27.57 *(3.25)*	29.13 *(3.62)*	22.18 *(3.38)*
Repetition (%)*+	33.49 *(3.65)*	49.54 *(4.32)*	27.79 *(3.11)*

*Post-hoc tests for repetition (Bonferroni), p <.05: * second different from first; ^ third different from first; + third different from second*

Regarding our cognitive load manipulation, we find a marginal effect of load on word-length [F (1, 82) = 3.31, p =.057, η_ρ^2 = .04], indicating that speakers who performed the task under high cognitive load articulated words faster than speakers in the low load condition. This is not surprising, given that we expect participants in the high load condition to "hurry" in their descriptions, in order to complete the task in time, which ultimately leads to a faster articulation. With respect to the influence of cognitive load on reduction, we find interactions between repetition and cognitive load with respect to the mean number of words contained in a referential phrase [F (2, 78) = 4.1, p <.05, η_ρ^2 = .09], the duration of the referential phrases [F (2, 78) = 5.8, p <.005, η_ρ^2 = .13], and the amount of common ground information [F (1, 39) = 8.5, p <.01, η_ρ^2 = .17]. Lack of space prohibits a further explanation of the these interaction effects, but generally they suggest that, even though reduction still takes place, the extent to which it occurs is mediated by the cognitive state of the speaker (see Fig. 3).

Reduction in gesture

With respect to gesture, we found that fewer referring expressions were accompanied by gestures in repeated mentions, as compared with initial mentions [F (1, 39) = 4.38, p <.05, η_ρ^2 = .1]. Cognitive load seems to influence the extent of this reduction, as evidenced by the interaction between repetition and cognitive load [F (1, 39) = 5.2, p <.05, η_ρ^2 = .11], with references produced by speakers under cognitive load being more often accompanied by gestures than the references produced by speakers in the low load condition (see Figure 3). The rest of the variables analyzed were not affected by cognitive load, although nearly all show an effect of repetition: the mean number of gestures per reference phrase [F (2, 78) = 7.91, p <.001, η_ρ^2 = .16], their duration [F (2, 78) = 73.8, p <.001, η_ρ^2 = .51], their size [F (2, 78) = 13, p <.001, η_ρ^2 = .25], and gestural repetition [F (2, 78) = 14.7, p <.001, η_ρ^2 = .27].

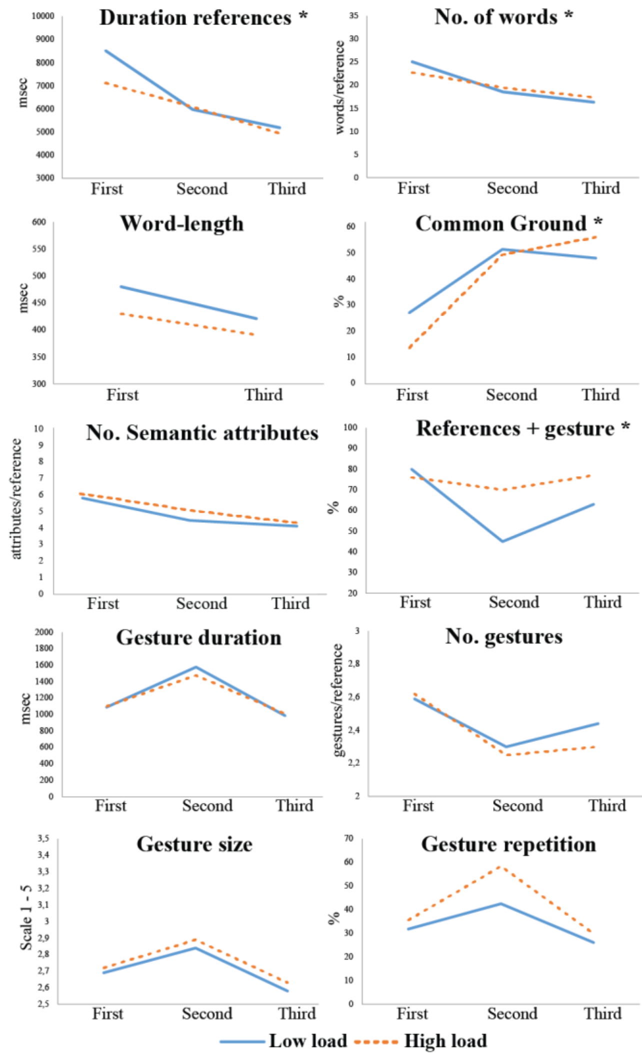

Figure 3: Overview of the results from Experiment I. Asterisks (*) indicate significant interactions between repetition and cognitive load.

Experiment II: perception

Method

In order to find out how communicative were the descriptions produced by speakers under the different experimental conditions from Experiment I, we conducted a perception test in which naïve participants had to attend to a number of referring expressions extracted from the production experiment footage (Experiment I), and match these expressions to the right target pieces, based on the principle that more communicative referring expressions would lead to higher percentages of correct answers.

Participants Ninety-seven Dutch students from Tilburg University (age M = 21.2 years; SD = 2.4, 73 female, 24 male), took part in this experiment in exchange for course credit. None of them had participated in Experiment I.

Stimuli The stimuli consisted of video and audio fragments containing referring expressions produced by the speakers from Experiment I. We selected one initial and one repeated (third) reference per speaker, and exported each of the fragments into three formats: a) audiovisual, b) video-only, and c) audio-only. We discarded data from nine participants who did not agree with their video recordings being shown to third parties, leaving us with referring expressions produced by 32 speakers. This created a total of 192 clips (64 clips per condition).

Procedure The perception test was administered online. When participants signed up to take part in the experiment they received a link to the online task, which randomly directed each new participant to one of the three experimental conditions (speech and gesture, gesture-only, or speech-only). Each participant attended to sixty-four clips containing referring expressions. The participants' task was to click on the picture that they thought corresponded to the speaker's description, being given four options (the target, plus three distractors of the same color set).

Statistical analyses We conducted a Repeated Measures ANOVA, with "repetition" (two levels) and "cognitive load" (two levels) as the within-subjects variables, and "condition" (three levels) as the between-subjects variable.

Results

As expected, participants who viewed clips in the "video-only" condition (*M* = .49, *SE* = .01) were less accurate at selecting the correct target than participants who attended to the clips in the "audiovisual" condition (*M* = .83, *SE* = .01) or in the "audio-only" condition (*M* = .81, *SE* = .01) [F (2, 94) = 173.6, p <.001, η_ρ^2 = .78]. Further, our results show that initial and repeated references were equally informative to participants, despite the decrease in the mean number of semantic attributes we found in our previous objective analyses. Interestingly, accuracy rates were higher when the participants viewed fragments produced by speakers under high load (*M* = .77, *SE* = .01) than when they viewed fragments produced by speakers in the low load condition (*M* = .64, *SE* = .009) [F (1, 94) = 233.04, p <.001, η_ρ^2 = .71].

Discussion

The present study explored the effect of cognitive load on the production of multi-modal referring expressions. Experiment I was able to replicate previous research, showing that repeated referring expressions are reduced with respect to initial ones in their speech (e.g., Bard et al., 2000; Howarth & Anderson, 2007). With respect to gestures, we observed reduction in the number of gestures produced by speakers in repeated references, in line with Hoetjes et al., (2011). Nevertheless, an interesting pattern arises for the mean duration of the gestures, their size and repetition, where we face an inverted "v-shaped" effect, with an increase in second mentions (increase in duration, in size, and in repetition), and a posterior decrease in third ones. We draw two conclusions from these patterns. First, we hypothesize that negative feedback from the addressees, or trouble in retrieving the correct piece during first trials, might have led speakers to produce longer and larger gestures in second trials. Further analyses of addressee feedback should clarify this. Second, these results show us that the reduction of gestural behaviours does not exactly parallel that of speech, suggesting that we are dealing with two independent, yet complementary processes (e.g., de Ruiter, 2000).

With respect to our main research question, we found reduction in the repeated references produced by speakers from both experimental conditions, suggesting that reduction takes place regardless of the degree of cognitive load experienced by the speaker. Nevertheless, as shown by the interactions in our data, cognitive load moderated the extent to which speakers reduced their utterances. Thus, for speakers under high load, reduction was less pronounced. This occurred for nearly all variables analyzed in speech, and for the percentage of referring expressions accompanied by iconic gestures. Nevertheless, at least for speech, only first references show differences across experimental conditions when there is an interaction effect (recall Figure 3), with speakers from both conditions reaching a similar end-level of reduction. Hence, we can conclude that both groups of speakers reduced their utterances to the same extent.

The question remains: is reduction mainly facilitated by speaker-internal or speaker-external processes? Some research posits that speakers reduce their utterances so that they are easier to process for their addressees, as a form of addressee adaptation (e.g., Fowler, 1988). Our results do not support this hypothesis, at least not if we consider this type of adaptation as being cognitively costly. Instead, our results are consistent with theories that view (articulatory) reduction as arising from generic language processes that are rather automatic (Dell & Brown, 1991; Bard et al., 2000). Thus, we contribute to these models by establishing that, not only articulatory, but also lexical and semantic reduction are part of the set of dialogical processes that take few cognitive resources to operate. We are nonetheless cautious about our results regarding the production of co-speech gestures: even though reduction in the amount of gestures was not influenced by cognitive load, speakers under high load tended to accompany their repeated references with gestures more often than speakers in the low load condition. This could imply that speakers under load may have benefitted from producing representational gestures (see, e.g., Kita, 2000).

Lastly, our common ground analyses show that, whereas the amount of shared information used by speakers increases with repetition for both groups, there are crucial differences in initial mentions, where participants in the low load condition provided their addressees with twice as much shared information than participants under high load. This shows a reluctance of speakers under load to adapt to their addressees in first mentions –consistent with Horton & Keysar (1996) and with Bard et al.'s (2000) dual model.

Experiment II aimed at complementing the results from our objective semantic analyses, by testing whether expressions produced in repeated references, or under cognitive load, would be any less (or any more) informative to naïve addressees. The results showed that referring expressions produced by speakers under cognitive load proved more communicative to the naïve judges, than utterances produced in low load. While this was an unexpected finding, we have two possible explanations. First, it can be that speakers in the low load condition tended to produce longer descriptions filled with hesitations, which might have made it more difficult for the listener to process them. Another explanation is that in the low load condition, speakers made more use of visually-based common ground with their addressees –e.g., by mentioning the spatial location of an item on the matcher's workspace, leading to descriptions equally rich in semantic attributes but not very communicative to naïve listeners without visual access to the workspace. Further analyses should help clarify this issue.

We conclude that, even though the reduction of repeated information might result into ease of processing for the addressee, this might not be the main motivation underlying it. We suggest that reduction could instead be a speaker-internal, load-lowering instrument that allows for a more efficient organization and packaging of thoughts.

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

Ariel, M. (1990). *Accessing noun phrase antecedents*, London/New York, Routledge.

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.

Bard, E. G., & Aylett, M. P. (2004). Referential Form, Word Duration, and Modeling the Listener in Spoken Dialogue. In John C. Trueswell and Michael K. Tanenhaus, eds. *Approaches to Studying World-Situated Language Use: Bridging the Language-as-Product and Language-as-Action Traditions*. Cambridge, MA: MIT Press.

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. (2000). The production of gesture and speech. In McNeill, D. (Ed.),*Language and Gesture* (pp. 248-311). Cambridge: Cambridge University Press.

Dell, G., & Brown, P. (1991). Mechanisms for listener adaptation in language production: Limiting the role of the "model of the listener". In D. J. Napoli & J. A. Kegl (Eds.), *Bridges between psychology and linguistics: A Swarthmore Festschrift for Lila Gleitman* (pp. 111–222). Hillsdale, NJ: Erlbaum.

Fowler, C. A. (1988). Differential shortening of repeated content words produced in various communicative contexts. *Language and Speech*, 31(4), 307-319.

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. (2011). GREEBLES Greeble greeb. On reduction in speech and gesture in repeated references. In L. Carlson, C. Hoelscher & T. F. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 3250-3255). Boston.

Holler, J., & Wilkin, K. (2011). An experimental investigation of how addressee feedback affects co-speech gestures accompanying speakers' responses. *Journal of Pragmatics, 43*, 3522–3536.

Holler, J., Tutton, M., & Wilkin, K. (2011). Co-speech gestures in the process of meaning coordination. In *Proceedings of the 2nd GESPIN - Gesture & Speech in Interaction Conference*, Bielefeld, 5-7 Sep 2011.

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.

Krauss, R. M., & Weinheimer, S. (1964) Changes in the length of reference phrases as a function of social interaction: A preliminary study, *Psychonomic Science, 1,* 113-114.

Lam, T. Q., & Watson, D. G. (2010). Repetition is easy: Why repeated referents have reduced prominence. *Memory and Cognition, 38*(8), 1137-1146.

Levy, R., & Jaeger, T. F. (2007). Speakers optimize information density through syntactic reduction. In *Proceedings of the twentieth annual conference on neural information processing systems* (pp. 849–856). Cambridge, MA: MIT Press.

McNeill, D. (1992). *Hand and mind. What gestures reveal about thought*. Chicago: University of Chicago Press.

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*. Austin, TX.

Roßnagel, C. (2000). Cognitive load and perspective-taking: Applying the automatic-controlled distinction to verbal communication. *European Journal of Social Psychology*, 30, 429 445.