Effects of cognitive load on the choice of referential form

Jorrig Vogels (j.vogels@uvt.nl)

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

Alfons Maes (maes@uvt.nl)
Tilburg center for Cognition and Communication, PO Box 90153
5000 LE Tilburg, the Netherlands

Abstract

We report on two experiments investigating the effect of an increased cognitive load for speakers on the choice of referring expressions. Speakers produced story continuations to addressees, in which they referred to characters that were either salient or non-salient in the discourse. In Experiment 1, all discourse information was shared. In Experiment 2, referents that were salient for the speaker were non-salient for the addressee, and vice versa. Cognitive load was manipulated by the presence or absence of a secondary task for the speaker. The results show that speakers under load are more likely to produce pronouns, at least when referring to characters that are less salient in their own discourse model. We take this finding as evidence that speakers under load are less able to determine that less salient characters should be referred to with more elaborate expressions, and hence are more likely to use economical expressions.

Keywords: cognitive load; reference; pronouns; language production; accessibility; perspective taking

Introduction

When speakers refer, they have to produce an expression that identifies the referent. Whether they choose a short expression such as 'she' or a more elaborate expression such as 'the girl' is generally assumed to be determined by the accessibility of the referent. According to the traditional view, speakers make assumptions about how accessible the referent is in the mind of their addressee (Ariel, 2001; Gundel, Hedberg, & Zacharski, 1993). The most important factor in determining this accessibility is assumed to be the salience of the referent in the discourse. For example, a referent that was the topic of the preceding sentence does not need an elaborate description to be reactivated in the memory of the addressee.

The choice of a referring expression may however also be influenced by speaker-internal constraints (e.g. Arnold, Bennetto, & Diehl, 2009; Arnold & Griffin, 2007). Speakers are not always monitoring the communicative needs of their addressees (e.g. Engelhardt, Bailey, & Ferreira, 2006; Fukumura & Van Gompel, 2012). One reason is that the language production system is constrained by the speaker's attention resources and working memory capacity. Since these resources are limited (Baddeley, 1986), speakers do not have unrestricted processing capacity to calculate the accessibility of referents for the addressee. In addition, it has been suggested that even when there is sufficient processing capacity, people do not always initially take the perspective of their conversational partners into account (e.g. Horton & Keysar, 1996).

Thus, it is not yet clear to what degree speakers base their choice of referential form on assumptions about how accessible the referent is for the addressee (e.g. Ariel, 2001; Gundel et al., 1993), and to what degree accessibility relates to how much attention the speaker herself has allocated to the referents in the discourse (e.g. Arnold & Griffin, 2007). One way to investigate this is to look at the effects of cognitive load on speakers' choice of referential form. Because taking into account the addressee's perspective and attending to discourse referents probably both take up memory resources (e.g. Horton & Keysar, 1996; Johnson-Laird, 1983), speakers that have an increased cognitive load (e.g. when they perform two tasks at the same time) may be less able to do this.

Depending on how speakers choose referring expressions, an increased cognitive load may affect this choice differently. Firstly, if speakers tailor references for their addressees, an increased cognitive load may make references more egocentric. That is, the choice of referring expression may be based on the speaker's own discourse model rather than on assumptions about their addressee's (e.g. Bard & Aylett, 2005; Fukumura & Van Gompel, 2012). Alternatively, speakers may use those expressions that are most economical for them to produce (e.g. pronouns; Hendriks, Koster, & Hoeks, 2013).

Secondly, if speakers choose referring expressions based on the degree to which they themselves attend to the referent, an increased cognitive load may reduce the accessibility of the referent, resulting in more elaborate expressions (e.g. Arnold & Griffin, 2007). Alternatively, speakers may have trouble keeping track of the salience of referents in their own discourse model, resulting in less consistent use of referring expressions (e.g. Arnold, 2010; Hendriks et al., 2013). That is, speakers experiencing high cognitive load may use fewer pronouns for referents that are salient in the discourse, and more pronouns for non-salient referents, compared to consistent use of pronouns for salient entities and full noun phrases for non-salient entities.

In this paper, we experimentally test these different possibilities, by investigating how an increased memory and attention load on the part of the speaker influences how speakers choose between attenuated expressions such as pronouns, and more elaborate expressions such as full noun phrases (full NPs). In two experiments, conducted in Dutch,

speakers told stories to an addressee, in which they referred to one of two characters, which was either salient or non-salient in the discourse. In one half of the experiment, they also performed a verbal memory task at the same time. In Experiment 1, all discourse information was shared between speaker and addressee. In Experiment 2, we dissociated the perspectives of the speaker and the addressee, such that whenever the referent was salient for the speaker, it was non-salient for the addressee, and vice versa.

Experiment 1

Method

Participants Sixty-four students (44 female; mean age 22.3 years) from Tilburg University participated in the experiment for course credit. Half of them acted as speakers, the others as addressees. All were native speakers of Dutch.

Materials The experimental items consisted of 16 pairs of photographs, taken from Vogels, Krahmer, and Maes (in press), accompanied by two introductory sentences and the onset of a third sentence. The first picture of a pair always showed one male and one female person. In the second picture, one of these persons performed an action, such as walking away or getting a glass of water. This person will be referred to as the target character, as participants were expected to refer to this character in their continuations. There were two versions of each picture pair; one in which the male person and one in which the female person performed the action. An example is shown in Figure 1.

The first sentence introduced both characters with indefinite NPs, which were the Dutch equivalents of either "a girl" and "a boy", or "a woman" and "a man". One of these was mentioned as the subject, the other in a prepositional phrase (PP), e.g., "A girl was arguing with a boy". The second sentence mentioned only the person that had been in the PP, e.g., "The boy got really annoyed", and was prerecorded by a female native speaker of Dutch. The onset of the third sentence was always "Subsequently...". In the discourse salient condition (condition A in Figure 1), the target character in the second picture was the subject of the second (i.e. directly preceding) sentence. In the discourse non-salient condition (condition B in Figure 1), the target character was the subject of the first sentence. Since all stories contained one male and one female character, pronoun references were never ambiguous.

In addition, 20 picture pairs served as fillers. To discourage participants from using only one type of expression throughout the experiment, some fillers showed either two male or two female characters (in which case pronouns were ambiguous) or only one character. In the accompanying sentences, some characters were given labels such as *een verkoopster* "a saleswoman" or *een Duitser* "a German", and sometimes the same character was the subject of both introductory sentences. An additional 4 items were included as practice items.

1. A girl was arguing with a boy. Subsequently...
2. The boy got really annoyed.

A: Discourse salient

1. A girl was arguing with a boy. Subsequently...
2. The boy got really annoyed.

B: Discourse non-salient

Figure 1: Example of a stimulus item in two conditions in Experiment 1. Context sentences are translations of the Dutch originals.

Procedure Two participants were randomly assigned to the role of speaker and addressee. The participant taking the role of speaker was seated at one end of a table, behind a laptop connected to a PST Serial Response Box. The participant taking the role of addressee was seated at the other end of the table, and was given a booklet containing all picture pairs and an answer sheet. The experiment was run on the laptop using E-Prime 2.0 software. The speaker's task was to complete the stories depicted by the picture pairs in such a way that the addressee could pick out the correct pair from the booklet.

Crucially, in one half of the experiment, the speaker performed a secondary task (cognitive load condition), while there was no secondary task in the other half (no cognitive load). In the no cognitive load condition, each trial started with the presentation of the item number, followed

by a fixation cross. Then, the first picture of a pair appeared on the left side of the screen, after which the first introductory sentence appeared below the picture in a red font. The speaker had 5 s to read this sentence aloud to the addressee, before the second sentence was presented over the computer speakers. Next, the second picture appeared to the right of the first picture, together with the text "Subsequently...". At this time, recording started, and the speaker had 6 s to complete the story. The addressee's task was to select the correct picture pair out of three options from the booklet and mark the correct answer on the sheet. In the experimental items, two of the three options differed only in which character performed the action, making correct reference crucial for the addressee to finish his task successfully. The addressee indicated to the speaker that the next trial could be started.

In the cognitive load condition, the appearance of the first picture was preceded by the word BAL or DAL,[1] which was presented on the screen for 1 s. The same happened at the end of the trial, after which the speaker indicated whether she saw the same or a different word by pressing the correct button on the response box. They did not receive feedback on their answers. We used a verbal rather than a visual secondary task (cf. Rosa & Arnold, 2011) to make sure that it would interfere with attending to discourse information rather than with visually attending to the characters in the pictures (e.g. Kellogg, Olive, & Piolat, 2007).

Speakers were instructed to pay attention to the auditorily presented sentences, since they would be asked about them after the experiment as an attention check. They were also encouraged to pay attention to the dual task by way of a prize offer for the participant with the fewest errors. To keep the speaker aware of the addressee's needs, the addressee was allowed to ask the speaker clarification questions if anything remained unclear, but only after the speaker had finished the story. The experiment was divided into two blocks, of which one contained the secondary task and the other did not, counterbalanced for order. Each block was preceded by two practice items.

Data coding Since the target referent was referred to only once in the majority of the cases, we only analyzed the first subject reference in each continuation. We excluded 1 case in which the first subject did not refer to the target referent and 1 plural reference (0.4%). The remaining 510 subject references were coded for the type of referring expression: either full NP or pronoun.

Design and statistical analyses Crossing the two factors referent salience and cognitive load resulted in a 2 (discourse salient, discourse non-salient) x 2 (cognitive load, no cognitive load) within-participants design. Participants were assigned to one of four lists, each of which contained one version of a given item. The items were presented in a

pseudo-random order, with at least one filler item between two consecutive experimental items.

We performed a logit mixed model analysis on the log odds for a pronoun (Jaeger, 2008). Referent salience and cognitive load were included as fixed factors; participants and items as random factors. The fixed factors were centered to reduce collinearity. We attempted to fit a model with a full random effect structure. In case the model did not converge, we excluded random slopes with the lowest variance (Barr, Levy, Scheepers, & Tily, 2013). From the first converging model, we subsequently excluded random slopes that did not significantly contribute to the model fit using log-likelihood ratio tests, with an α-level of .20 (Barr et al., 2013). Only the final model will be reported.

Results and Discussion

Figure 2 shows the proportion of pronoun references to the target character by referent salience and cognitive load condition. The final logit mixed model included random intercepts for participants and items, and a by-participant random slope for referent salience. We found a main effect of referent salience: pronouns were more frequent when the referent was discourse salient (54.1%) than when it was not (11.0%), $\beta = 6.24$; SE = 0.88; $p < .001$. There was also a significant main effect of cognitive load: more pronouns were used when speakers performed a secondary task (34.5%) than when they did not (30.6%), $\beta = 0.76$; SE = 0.39; $p < .05$. These effects were qualified by a marginally significant interaction, $\beta = -1.34$; SE = 0.77; $p = .08$.
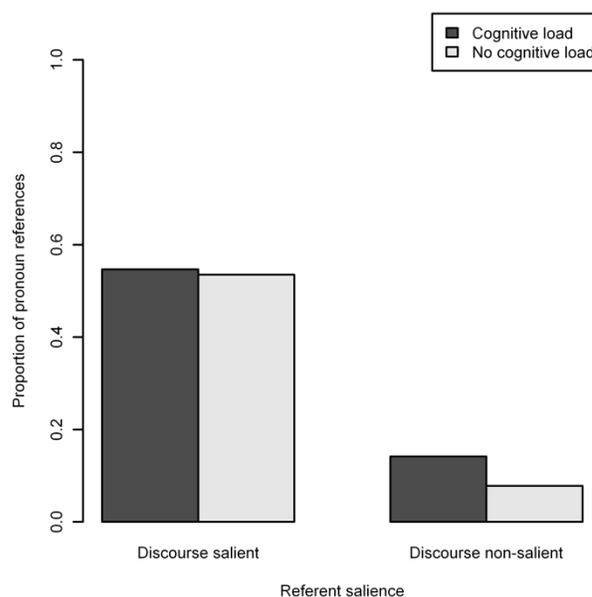

Figure 2: Proportion of pronoun references to the target character in the four conditions in Experiment 1.

To arrive at the pairwise comparisons for the interaction effect, we built separate models for the two levels of referent salience. The final model for the discourse salient condition included only by-participant and by-item random

---

[1] Meaning "ball", and "valley", respectively. These stimuli were adopted from Goudbeek and Krahmer (2011).

intercepts, and the model for the discourse non-salient condition included only a by-participant random intercept. The effect of cognitive load was not significant in the discourse salient condition, $\beta = 0.11$; SE = 0.38; $p = .78$, but significant in the discourse non-salient condition, with pronouns being more frequent in the cognitive load condition (14.2%) than in the no cognitive load condition (7.8%), $\beta = 1.55$; SE = 0.67; $p < .05$.

These results indicate that adding cognitive load increases the probability of pronoun use. This effect only seems to be present when the referent is not salient in the discourse. This suggests that an increased cognitive load does not reduce the accessibility of referents in the speaker's own discourse model. To test whether cognitive load makes it more difficult to take into account the perspective of the addressee, Experiment 2 investigated the effect of cognitive load when the speaker's and addressee's perspectives were dissociated. This was done by presenting the second context sentence only to the speaker, over headphones (cf. Fukumura & Van Gompel, 2012). In this way, the referent that was most salient for the speaker was not salient for the addressee (speaker-salient condition), since the addressee did not hear the sentence in which that referent was in subject position. Conversely, the referent that was less salient for the speaker was most salient for the addressee (addressee-salient condition), since it was the most recent subject for the addressee. We predicted that if speakers are worse at perspective taking under load, they should be less likely to take into account the accessibility of the referent for the addressee when performing a secondary task. That is, they should be more likely to choose referring expressions based on their own perspective.

Experiment 2

Method

Participants Sixty-four students (47 female; mean age 20.2 years) from Tilburg University participated in the experiment for course credit. Half of them acted as speakers, the others as addressees. All were native speakers of Dutch. None of them participated in Experiment 1.

Materials We used the same materials as in Experiment 1.

Procedure The procedure was identical to that of Experiment 1, except that the second context sentence was now presented only to the speaker, over headphones. Speakers were explicitly told that this sentence could not be heard by their addressee, but that they had to pay attention to it nonetheless, since they would be asked about it after the experiment.

Data coding The data coding procedure was the same as in Experiment 1. We excluded 34 cases in which the first subject did not refer to the target referent. In addition, we excluded 7 plural references, 3 indefinite references, 1 case

in which the sentence presented over the headphones was repeated literally, and 1 missing case (9.0%). The remaining 466 subject references were coded for the type of referring expression: either full NP or pronoun.

Design and statistical analyses Crossing the two factors referent salience and cognitive load resulted in a 2 (speaker-salient, addressee-salient) x 2 (cognitive load, no cognitive load) within-participants design. Statistical analysis of the data was done in the same way as in Experiment 1.

Results and Discussion

Figure 3 shows the proportion of pronoun references to the target character by referent salience and cognitive load condition. The final logit mixed model included random intercepts for participants and items, as well as by-participant random slopes for referent salience and cognitive load. We found a main effect of referent salience: pronouns were more frequent when the referent was discourse salient only for the speaker (23.6%) than when it was salient only for the addressee (8.3%), $\beta = 2.25$; SE = 0.85; $p < .01$. There was also a significant main effect of cognitive load, with slightly more pronouns in the cognitive load condition (17.2%) than in the no cognitive load condition (15.8%), $\beta = 1.37$; SE = 0.56; $p < .05$. However, these effects were qualified by a significant interaction between referent salience and cognitive load, $\beta = -2.76$; SE = 0.95; $p < .01$.
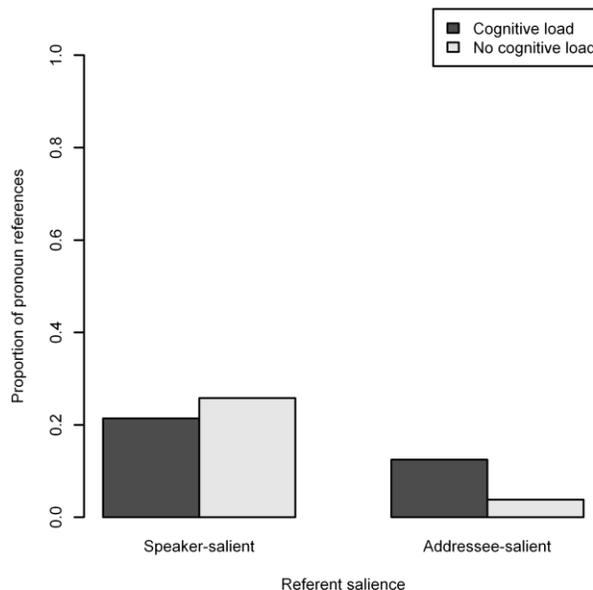

Figure 3: Proportion of pronoun references to the target character in the four conditions in Experiment 2.

We investigated this interaction further by building separate models for the two levels of referent salience. The final model for the speaker-salient condition included a by-participant random slope for cognitive load; the model for the addressee-salient condition included only by-participant and by-item random intercepts. The effect of cognitive load

was not significant in the speaker-salient model, $\beta = 0.77$; SE = 0.61; $p = .21$, but marginally so in the addressee-salient model, with pronouns being more frequent in the cognitive load condition (12.5%) than in the no cognitive load condition (3.8%), $\beta = 1.35$; SE = 0.77; $p = .08$.

These results indicate that, as in Experiment 1, the presence of the secondary task increased the likelihood of pronoun use, at least when the referent was not salient for the speaker (addressee-salient condition). Comparing the results of Experiment 1 and Experiment 2 indeed suggests that the effect of cognitive load does not differ across the two experiments. A logit mixed model with referent salience (salient for the speaker, not salient for the speaker), cognitive load (cognitive load, no cognitive load), and experiment (shared context, privileged context) as fixed factors, and participants and items as random factors (including random slopes for referent salience) showed that there was no significant interaction between cognitive load and experiment, $\beta = -0.24$; SE = 0.61; $p = .69$ and no three-way interaction, $\beta = -0.58$; SE = 1.25; $p = .64$.

This finding does not support the hypothesis that speakers are more likely to fall back on their own discourse model when they experience an increased cognitive load. If that were the case, pronouns should have become less frequent under load when the referent was not salient for the speaker but salient for the addressee, and more frequent in the reverse situation. In fact, speakers did not seem to take into account the addressee's perspective even when they were not performing a secondary task (i.e. they were still more likely to use pronouns when the referent was salient for them). This is in line with Fukumura and Van Gompel (2012), who found that speakers tend to follow their own discourse model when there is privileged information. Pronoun use did decrease in the speaker-salient condition compared to Experiment 1, suggesting some (minimal) audience design (Galati & Brennan, 2010), but this was not affected by cognitive load. Thus, the effect of cognitive load seems to be independent of perspective taking.

General Discussion

In both experiments, speakers were more likely to produce pronouns in their continuations when they were performing a secondary task at the same time, at least when referring to a non-salient character. This finding does not support the hypothesis that an increased cognitive load reduces the accessibility of referents in the speaker's own discourse model, since that would have resulted in an increase of full NPs. This suggests that although accessibility may be related to attention, it may not hold generally that attentional load leads to more elaborate referring expressions. In addition, our results do not support the hypothesis that an increased cognitive load affects the speaker's ability to take the perspective of the addressee, since speakers did not appear to calculate the referent's accessibility for the addressee even when they were not under load. Rather, they seemed to employ some kind of audience design by increasing the use of full NPs when there was privileged

information (cf. Fukumura & Van Gompel, 2012), which was not harmed by the execution of a secondary task.

Based on these results, we conjecture that speakers tailor their referring expressions for the addressee by basing their choices on the salience of referents in their own discourse model. This model built from the speaker's perspective may serve as a proxy for that of the addressee (e.g. Bard & Aylett, 2005), and can therefore still be regarded as audience design. Thus, even though speakers do not seem to specifically keep track of the accessibility of the referent for the addressee, the production of more costly referring expressions such as full NPs may be inherently oriented towards an addressee (cf. Hendriks et al., 2013). Under load, however, speakers have fewer memory resources available to determine that referents that are not salient in the discourse should be referred to with more elaborate expressions. Therefore, they are more likely to resort to using expressions that are most economical for them, i.e. pronouns (e.g. Burzio, 1998). Since this preference coincides with the preference to use pronouns for discourse salient referents, cognitive load does not make a difference when the referent is salient for the speaker.

As noted above, in addition to the use of full NPs for non-salient referents, another type of audience design, emerging from the comparison of the two experiments, is that speakers are more likely to use full NPs as soon as it is clear that the preceding discourse is not fully in common ground with their addressee. This may be evidence for a minimal, one-bit model of audience design (Galati & Brennan, 2010): speakers use more specific referring expressions as soon as they are aware that not all information is shared, but irrespective of the actual accessibility of the referent for the addressee. This is in line with Fukumura and Van Gompel (2012), who found that while speakers were not taking into account their addressee's perspective in choosing referring expressions when the two perspectives were dissociated, they used slightly more pronouns in a condition in which all information was shared (37% vs. 33%), independently of whether the referent was salient or not. This suggests that speakers use more elaborate expressions when there is privileged information, even though they might run the risk of being overly specific.

One reason why speakers did not make the extra effort to calculate the accessibility of the referent in the addressee's discourse model may be that in the current experiment, as well as in Fukumura and Van Gompel's, references were never ambiguous, since the two characters always had a different gender. Therefore, not taking into account the addressee's perspective would probably not result in interpretation errors. However, in case not taking into account the addressee's perspective *would* lead to interpretation errors, speakers may be more likely to base their choice of referring expressions on the discourse model of their addressee (e.g. Fukumura & Van Gompel, 2012; Horton & Keysar, 1996). In that case, increased cognitive load might make this perspective taking more difficult, and cause speakers to fall back on their own discourse model.

It is important to note that the effect of cognitive load on referential choices may depend on the specific task and on what part of cognition is actually loaded. In a dual-task setup, the nature of the secondary task may make a difference (e.g. whether it is visual or verbal; Baddeley, 1986; Kellogg et al., 2007). In addition, manipulations such as varying the number of referential competitors (e.g. Arnold and Griffin, 2007) might lead to a different kind of cognitive load than divided attention over multiple tasks. It is also conceivable that our artificial dissociation of perspective using privileged information presented over headphones also caused some increase in cognitive load. Finally, our results suggest that there was quite some individual variation as to how speakers' referring expressions were affected by the dual task. Although our cognitive load manipulation had an impact on referential choices, the secondary task appeared to be relatively easy. Informal inspection of the data suggested that participants who found the task difficult showed the clearest effects of cognitive load. These issues need further research.

Our results are potentially relevant for computational models of referring expression generation, as they suggest that referring expressions are not always in accordance with the referent's discourse salience (cf. Krahmer & Theune, 2002). Psychologically plausible models should incorporate that sometimes pronouns are used when there are insufficient resources to take salience into account.

References

Ariel, M. (2001). Accessibility theory: an overview. In T. Sanders, J. Schilperoord & W. Spooren (Eds.), *Text Representation: Linguistic and psycholinguistic aspects* (pp. 29-87). Amsterdam/Philadelphia: John Benjamins.

Arnold, J. E. (2010). How Speakers Refer: The Role of Accessibility. *Language and Linguistics Compass, 4*(4), 187-203.

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

Arnold, J. E., Bennetto, L., & Diehl, J. J. (2009). Reference production in young speakers with and without autism: Effects of discourse status and processing constraints. *Cognition, 110*, 131-146.

Baddeley, A. (1986). *Working Memory*. Oxford: Clarendon Press.

Bard, E. G., & Aylett, M. P. (2005). Referential Form, Word Duration, and Modeling the Listener in Spoken Dialogue. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to Studying World-Situated Language Use: Bridging the Language-as-Product and Language-as-Action Traditions* (pp. 173-191). Cambridge, MA: MIT Press.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Burzio, L. (1998). Anaphora and Soft Constraints. In P. B. D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky (Eds.), *Is the Best Good Enough?: Optimality and Competition in Syntax*. Boston: MIT.

Engelhardt, P. E., Bailey, K. G. D., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of Quantity? *Journal of Memory and Language, 54*, 554-573.

Fukumura, K., & Van Gompel, R. (2012). Producing Pronouns and Definite Noun Phrases: Do Speakers Use the Addressee's Discourse Model? *Cognitive Science, 36*(7), 1289-1311.

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language, 62*(1), 35-51.

Goudbeek, M., & Krahmer, E. (2011). Referring under load: Disentangling preference-based and alignment-based content selection processes in referring expression generation *Proceedings of PRE-Cogsci: Bridging the gap between computational, empirical and theoretical approaches to reference*. Boston.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69*, 274-307.

Hendriks, P., Koster, C., & Hoeks, J. C. J. (2013). Referential choice across the lifespan: why children and elderly adults produce ambiguous pronouns. *Language and Cognitive Processes*.

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

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434-446.

Johnson-Laird, P. N. (1983). *Mental models: towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Kellogg, R. T., Olive, T., & Piolat, A. (2007). Verbal, visual, and spatial working memory in written language production. *Acta Psychologica, 124*, 382–397.

Krahmer, E., & Theune, M. (2002). Efficient Context-Sensitive Generation of Referring Expressions. In K. v. Deemter & R. Kibble (Eds.), *Information sharing: Reference and Presupposition in Language Generation and Interpretation* (pp. 223-264). Stanford, CA: CSLI Publications.

Rosa, E. C., & Arnold, J. E. (2011). The Role of Attention in Choice of Referring Expression *Proceedings of PRE-Cogsci: Bridging the gap between computational, empirical and theoretical approaches to reference*. Boston.

Vogels, J., Krahmer, E., & Maes, A. (in press). Who is where referred to how, and why? The influence of visual saliency on referent accessibility in spoken language production. *Language and Cognitive Processes*.