Effects of Cognitive Effort on the Resolution of Overspecified Descriptions

Ivandré Paraboni*
USP / EACH

Alex Gwo Jen Lan*
USP / EACH

Matheus Mendes de Sant'Ana*
USP / EACH

Flávio Luiz Coutinho*
USP / EACH

Studies in referring expression generation (REG) have shown different effects of referential overspecification on the resolution of certain descriptions. To further investigate effects of this kind, this article reports two eye-tracking experiments that measure the time required to recognize target objects based on different kinds of information. Results suggest that referential overspecification may be either helpful or detrimental to identification depending on the kind of information that is actually overspecified, an insight that may be useful for the design of more informed hearer-oriented REG algorithms.

1. Introduction

In natural language generation, referring expression generation (REG) is the computational task of producing descriptions that enable the hearer to identify a target object (Dale and Reiter 1995). REG may be divided into two fairly independent subtasks: content selection (deciding what to say) and surface realization (deciding how to say it). In this article we will focus on the content selection task of definite descriptions, as in "the man wearing a dark suit," "the young man on the left," and so forth.

Central to the present discussion is the issue of referential overspecification (Pechmann 1989). The inclusion of more information than required for disambiguation (e.g., to produce descriptions as in "the bird *on top of the mailbox*" in a context in which there is only one visible bird) is common in language use, and it is known to affect reference resolution in a number of ways. In particular, studies such as in Arts et al. (2011) have shown that overspecified descriptions may help identification,

---

∗ Av. Arlindo Bettio, 1000 - São Paulo, Brazil. E-mail: `ivandre@usp.br`.

whereas studies such as in Engelhardt, Demiral, and Ferreira (2011) have shown that adding certain modifiers in this way may actually slow down comprehension.

Knowing whether a particular piece of information may help or impair identification is crucial for the design of hearer-oriented REG algorithms. Assuming that we would like to generate descriptions that are easy to identify, this knowledge may shed light on long-standing issues in REG such as the question of when to use **relational properties** for the purpose of identification (e.g., as in the "bird *on top of the mailbox*"). On the one hand, algorithms such as in Dale and Haddock (1991) make use of relations as a last resort, that is, only when it is not possible to produce a uniquely identifying atomic description. On the other hand, studies such as Viethen and Dale (2011) suggest that using relations may be common even when relations are not required for disambiguation. Clearly, one possible way of deciding when to use a relational property is by assessing its impact on identification.

To investigate when referential overspecification may help or impair identification, we report on two eye-tracking experiments that measure the time spent examining objects in a visual context based on different kinds of information. Our findings suggest that easily recognizable properties may help identification, whereas other properties may have the opposite effect. These results are consistent with previous work in the field, and pave the way for the design of more informed hearer-oriented REG algorithms.

2. Related Work

Human speakers often include redundant information when referring to a target object, and existing approaches to REG have long attempted to mimic this behavior (Dale and Reiter 1995). Generally speaking, however, these studies do not address the question of whether referential overspecification may be helpful or not from the hearer's perspective. Exceptions are briefly discussed here.

A number of studies in REG and related fields have found that overspecified descriptions may lead to *faster* identification. The work in Arts et al. (2011) compares identification times required by minimally distinguishing descriptions (e.g., "the button") and overspecified alternatives (e.g., "the round white button"). Results show that identification is faster (or at least not slower) when additional information is presented. The study also considered the use of overspecified spatial relations, and found that including information about vertical position (e.g., above, below) also leads to faster identification times.

Similarly, the work in Paraboni and van Deemter (2014) considers the use of overspecified spatial relations to facilitate search in large or structurally complex spatial domains. The study focuses on a number of situations in which minimally distinguishing descriptions may lead to confusion or even misidentification, as in "the button behind a chair" in a context in which there are two chairs, but the intended one is not the nearest chair from a hearer's perspective. In situations of this kind, it is argued that the use of overspecified spatial relations, as in "the button behind a chair, *on the left side*" not only help, but may be actually *required* for successful identification.

Other studies, by contrast, have found that referential overspecification may be *detrimental* to identification. In Engelhardt, Baileyand, and Ferreira (2006), eye tracking and ERP studies showed that hearers require additional time to execute instructions that contain an overspecified prepositional phrase modifier as in "put the apple *on the towel* in the box" in a context containing one apple on a towel and a second empty towel.

Referential overspecification in this case was found to lead to temporary confusion about where the apple should be placed. Similarly, the work in Engelhardt, Demiral, and Ferreira (2011) assessed how quickly participants could orient attention to an object upon hearing descriptions as in "the red square" in a context in which the reference to color could be either necessary or not. Results showed that referential overspecification leads to longer reaction times.

3. Current Work

We now report on two eye-tracking experiments that measure the time spent examining objects in a visual context based on different kinds of information. The first experiment (Section 3.1) focuses on the use of overspecified relational properties, and the second (Section 3.2) focuses on the use of atomic properties. In both cases, we would like to show that referential overspecification may be either helpful or detrimental, depending on the effort required to recognize the additional information.

3.1 Experiment 1: Overspecification Using Relational Properties

Consider a simple domain conveying seven- and eight-pointed stars with alphanumeric labels as in Figure 1. In this domain, we will focus on situations of reference in which a target description (e.g., "the letter") may be overspecified with the addition of a relational property (e.g., "the letter *inside an eight-pointed star*").

In this domain, we assume that recognizing an object shape (i.e., determining whether a star has seven or eight points) requires more time than recognizing a label type (i.e., determining whether a star contains a letter or number). Thus, we will say that object shapes are comparatively "difficult" to recognize, and that label types are comparatively "easy." This basic assumption, although plausible in the current context, will be verified as part of the subsequent data analysis.

We consider four experimental conditions in Table 1. Conditions *min.D* and *min.E* make use of atomic descriptions that may be either difficult (e.g., "the eight-pointed star") or easy ("the letter") to recognize. These will be compared to alternatives in which a second property is overspecified, and which may in principle be either difficult or easy as well. In the original (Portuguese) descriptions considered in the experiment, the added information is to be interpreted as being restrictive, that is, readers will assume that the information is in principle required for disambiguation.

In this setting, our goals are twofold: We would like to show that overspecified relational properties that are easy to recognize may help identification, and that those that are difficult to recognize may have the opposite effect. To this end, we designed
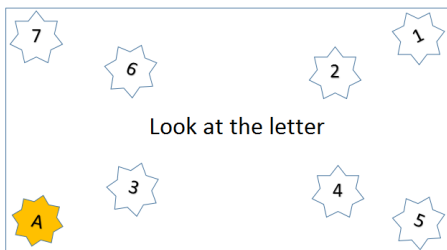

Figure 1
A stimulus image and example instruction for Experiment 1. The target is presently highlighted for ease of discussion.

Table 1
Conditions for Experiment 1.

Condition	Translated example	Portuguese original description
min.D	the 8-pointed star	a estrela de 8 pontas
overspec.DE	the 8-pointed star *that contains a letter*	a estrela de 8 pontas *que contém uma letra*
min.E	the letter	a letra
overspec.ED	the letter *inside an 8-pointed star*	a letra *dentro de uma estrela de 8 pontas*

an experiment in which participants were required to identify target objects in stimulus images as in Figure 1. In doing so, we measured the time spent examining every individual object on screen with the aid of a gaze-tracking device. In order to minimize the effects of task awareness, however, our focus is on the inspection of distractor objects, and not on the target itself.[1] Our research hypotheses are as follows:

h1a: *Adding an easily recognizable relational property to a minimally distinguishing description will decrease mean inspection times.*

h1b: *Adding a relational property that is difficult to recognize to a minimally distinguishing description will increase mean inspection times.*

Hypothesis *h1a* predicts that referential overspecification facilitates identification along the lines of Arts et al. (2011) and Paraboni and van Deemter (2014). This will be tested by comparing mean inspection times required by minimally distinguishing descriptions conveying a property that is difficult to recognize (condition *min.D*, as in "the eight-pointed star") with overspecified alternatives using a property that is easier to recognize (*overspec.DE*, as in "the eight-pointed star *containing a letter*"). Inspection times are expected to decrease from *min.D* to *overspec.DE*.

Hypothesis *h1b* predicts that referential overspecification impairs identification along the lines of Engelhardt, Demiral, and Ferreira (2011) and Engelhardt, Baileyand, and Ferreira (2006). This will be tested by comparing inspection times required by minimally distinguishing descriptions conveying only the property that is easy to recognize (condition *min.E*, as in "the letter") with the overspecified alternatives using the property that is more difficult to recognize (*overspec.ED*, as in "the letter *inside an eight-pointed star*"). Inspection times are expected to increase from *min.E* to *overspec.ED*.

In both cases, we notice that our focus is on the comparison between short and long descriptions with restrictive reading, and we do not compare the two overspecified alternatives directly. This contrasts with the work in Clarke, Elsner, and Rohde (2015), for example, which investigates the relation between order of mention and visual salience.

Subjects 25 undergraduate students acting as volunteers. Participants were on average 20.4 years old (min = 18 and max = 23), predominantly male (19, or 76%), and had normal or corrected vision. All participants were native Portuguese speakers.

Procedure Each trial was run individually, and included practice time to familiarize subjects with the task. Participants were required to identify the target objects

---

1 For instance, target identification had to be fully acknowledged by pressing the space bar, a decision that may affect its overall visualization time in a way that does not occur in the case of distractors.

Table 2
Experiment 1: Inspection time variation obtained by adding easy / difficult properties.

Cognitive Effort	Estimate	Std. Error	t value	p value
easy	–136.07	37.24	–3.654	<0.001
difficult	262.86	49.19	5.343	<0.001

mentioned on each scene and press the space bar to acknowledge this. No mention was made of the distractor objects, which were the true focus of the experiment. Stimulus images were presented individually in random order, and included the instruction conveying the description of the intended target, presented (in Portuguese) in the central portion of the screen, as in "Look at the eight-pointed star." Inspection times were computed by considering all visits to distractor objects on screen, from the moment in which the scene was displayed until identification. When the participant's gaze was not upon any of the eight objects on screen, inspection times were not computed.

Materials   The experiment made use of purpose-built software to present the stimuli and to compute inspection times of screen objects. We used a set of 24 images similar to those in Figure 1 representing 8 experimental items and 16 fillers. All images showed eight star objects with a description in the center.[2]

As a means to increase the chances that at least one distractor object would be visited before reaching the target, in the experimental items the target object is always placed in one of the screen corners. For that reason, filler items were included so as to place targets in the inner screen positions as well. The angle of inclination of each object was altered so as to make the recognition of the star shape more difficult, and it has no particular meaning for the experiment. Also as a means to increase the chances of visiting at least one distractor object during search, each experimental condition was tested twice by alternating label types (letters/numbers) and shapes (seven- or eight-pointed), and also by varying object positions.

Results   Table 2 shows the mean gain in distractor inspection times for the two over-specification strategies over the minimally distinguishing alternatives. All p values were obtained by using the Satterthwaite approximation.

Before discussing our main hypotheses, we will briefly examine the basic assumption upon which the experiment was built, namely, the assumption that recognizing a star shape (seven- or eight-pointed) is more difficult than recognizing a label type (letter/number). We found that referring to an object shape (*min.D*) demands longer identification times than referring to a label type (*min.E*). The difference is significant, according to a one-way ANOVA test ($F(1,48) = 16.3442$, MSE = 22133.65 $p < 0.001$). We also compared the time required by descriptions that refer to letters and numbers, and by descriptions that refer to seven- and eight-pointed objects. In both cases, there was no significant difference in using either alternative.

For our main hypotheses, we report results from a linear mixed-effects analysis of the relationship between distractor recognition times and cognitive effort. As fixed effect, we entered cognitive effort (easy / difficult) into the model. As random effects

---

2 Although we could have used simple object pairs alone, the use of more complex scenes will support further investigation on the choice of search paths (to be discussed in a subsequent paper).

we had intercepts for subjects and items, and also by-subject and by-item random slopes for the effects of cognitive effort.

Regarding hypothesis *h1a*, we notice that adding an easily recognizable property decreases inspection times. This offers support to hypothesis *h1a*. Regarding *h1b*, the opposite effect is observed: Adding a property that is more difficult to recognize increases inspection times. A likelihood ratio test of the full model against the model without this effect shows that the difference is significant ($\chi^2(1) = 4.9366$, p $= 0.02629$).

Discussion Because readers do not know in advance whether a piece of information is overspecified or not, they are forced to take the entire description into account when searching for its target. Thus, if the redundant information happens to be easily recognizable (e.g., an object label), the additional property may facilitate identification. This outcome is consistent with Arts et al. (2011), Paraboni and van Deemter (2014), and others. On the other hand, if the redundant information turns out to be difficult to recognize (e.g., star shape), then referential overspecification will have the opposite effect. This outcome is consistent with studies such as Engelhardt, Demiral, and Ferreira (2011).

3.2 Experiment 2: Overspecification Using Atomic Properties

In the domain considered in Experiment 1, the difference between so-called easy and difficult properties was self-explanatory. In our second experiment, we would like to provide further evidence that referential overspecification may help or impair identification in more subtle situations by considering a visual domain containing character pairs of letters and numbers, as in Figure 2, and by focusing on situations in which the landmark portion of a relational description may either refer to a more general class of objects (e.g., "the even number followed by *a letter*") or to a subcategory (e.g., "the even number followed by *a consonant*").

Analogously to the recognition of star shapes and labels in Experiment 1, we assume that recognizing whether a character is a letter is easier than recognizing whether it is a consonant, and we will make use of three experimental conditions *min*, *overspec.G*, and *overspec.S* in Table 3. These conditions differ from each other only in the kind of referring expression under consideration. The *min* condition makes use of the minimally distinguishing atomic description "the even number." In both *overspec* conditions, this same atomic description is overspecified by adding a reference to a landmark object (i.e., "the adjacent character"). Condition *overspec.G* refers to the more general property
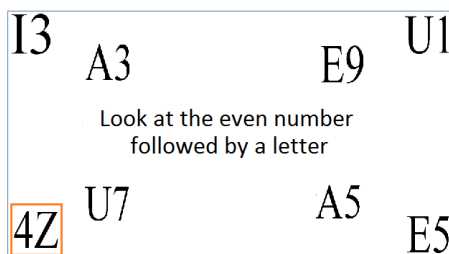

Figure 2
A stimulus image and example instruction for Experiment 2. The target is presently highlighted for ease of discussion.

Table 3
Conditions for Experiment 2.

Condition	Translated example	Portuguese original description
min	the even number	o número par
overspec.G	the even number *followed by a letter*	o número par *seguido de uma letra*
overspec.S	the even number *followed by a consonant*	o número par *seguido de uma consoante*

"letter," whereas *overspec.S* refers to the more specific property "consonant." As in Experiment 1, the overspecified alternatives (in Portuguese) have a restrictive reading.

An identification task similar to Experiment 1 was carried out. Our research hypotheses are as follows:

h2a: *Adding a more general atomic property to a minimally distinguishing description will decrease mean inspection times.*

h2b: *Adding a more specific atomic property to a minimally distinguishing description will increase mean inspection times.*

Hypothesis *h2a* predicts that overspecification using the more general property "letter" facilitates identification just like any property that is easily recognizable (cf. Experiment 1). This will be tested by comparing inspection times required by minimally distinguishing descriptions (*min*, as in "the even number") with the over-specified alternative using a more general property (*overspec.G*, as in "the even number *followed by a letter*"). Inspection times are expected to decrease from *min* to *overspec.G*.

Hypothesis *h2b* predicts that referential overspecification using the more specific property "consonant" will impair identification just like those properties that are diffi-cult to recognize (cf. Exp.1). This will be tested by comparing inspection times required by minimally distinguishing descriptions (*min*, as in "the even number") with the overspecified alternative using a more specific property (*overspec.S*, as in "the even number *followed by a consonant*"). Inspection times are expected to increase from *min* to *overspec.S*.

Subjects   25 undergraduate students who responded to an invitation to act as volun-teers. Participants were on average 20.4 years old (min = 18 and max = 25) and had normal or corrected vision. All participants were male native speakers of Portuguese.

Procedure   Same as in Experiment 1.

Materials   Stimuli consisted of 20 images similar to those in Figure 2 representing 6 experimental items and 14 images regarded as fillers for our current purposes. All images showed eight character pairs with a description in the center. Once again, ex-perimental items always placed the target object in one of the screen corners, and fillers were designed so as to place the target in the four inner positions as well. Further scene variation was implemented by alternating letters and numbers, and also by changing the object positions across images.

Results   Table 4 shows the mean gain in distractor inspection times for the two overspecification strategies over the *min* alternative, with p values obtained from the

Table 4
Experiment 2: Inspection time variation obtained by adding general / specific properties.

Cognitive Effort	Estimate	Std. Error	t value	p value
general	−72.22	26.19	−2.757	0.01062
specific	121.73	30.08	4.047	0.00022

Satterthwaite approximation. We report results from a linear mixed-effects analysis of the relationship between distractor recognition times and specificity, having specificity (general / specific) as a fixed effect, intercepts for subjects and items, and also by-subject and by-item random slopes for the effects of specificity.

Regarding hypothesis *h2a*, we notice that adding a more general property decreases mean inspection times. Regarding *h2b*, we notice that adding a more specific property has the opposite effect. A likelihood ratio test of the full model against the model without this effect shows that the difference is significant ($\chi^2(1) = 13.28$, p = 0.009984).

Discussion   Once again, referential overspecification may be either helpful or detrimental, depending on the effort (presently modeled by examples of general/specific properties) required to process redundant information marked as restrictive. Adding an easily recognizable property ("letter") helps, whereas adding a more "difficult" property ("consonant") impairs identification. This outcome is consistent with Experiment 1, but it is now observed in the use of a redundant atomic property (as opposed to using a redundant relational property, cf. Section 3.1).

4. Final Remarks

This article addressed the issue of how the effort involved in the recognition of overspecified information may affect identification. Our findings suggest that easily recognizable properties may facilitate identification, whereas properties that are more difficult to recognize may have the opposite effect. These observations were shown to hold for relational and atomic properties alike, and are in principle consistent with existing studies that argue that referential overspecification helps identification, and also with those that claim the opposite.

From a computational perspective, these insights are potentially relevant in a number of ways. In particular, we notice that inspection times may provide a suitable account of salience for hearer-oriented REG. Knowing the amount of time required to recognize different kinds of property may guide attribute selection in incremental-like algorithms (Dale and Reiter 1995), for example, by considering properties that require less cognitive effort first. Similarly, the effort required to recognize a relational property may guide the decision of whether to use a relation or not, which is a longstanding issue in the REG field (Dale and Haddock 1991; Viethen and Dale 2011). The design of a REG algorithm along these lines—which would require measuring cognitive effort in advance—is left as future work.

We note also that our study was carried out by taking a hearer-oriented perspective, that is, by assuming that we would like to generate descriptions that facilitate identification. This raises the question of whether speakers may actually favor the use of easy properties in this way, or how often they do so. A speaker-oriented study of this kind is also left as future work.

References

Arts, A., A. Maes, L. G. M. Noordman, and C. Jansen. 2011. Overspecification facilitates object identification. *Journal of Pragmatics*, 43(1):361–374.

Clarke, Alasdair D. F., Micha Elsner, and Hannah Rohde. 2015. Giving good directions: Order of mention reflects visual salience. *Frontiers in Psychology*, 6:1–10.

Dale, R. and N. J. Haddock. 1991. Content determination in the generation of referring expressions. *Computational Intelligence*, 7(4):252–265.

Dale, R. and E. Reiter. 1995. Computational interpretations of the Gricean maxims in the generation of referring expressions. *Cognitive Science*, 19(2):233–263.

Engelhardt, P. E., K. Baileyand, and F. Ferreira. 2006. Do speakers and listeners observe the Gricean maxim of quantity? *Journal of Memory and Language*, 54(4):554–573.

Engelhardt, P. E., S. B. Demiral, and Fernanda Ferreira. 2011. Over-specified referring expressions impair comprehension: An ERP study. *Brain and Cognition*, 77(2):304–314.

Paraboni, Ivandré and Kees van Deemter. 2014. Reference and the facilitation of search in spatial domains. *Language, Cognition and Neuroscience*, 29(8):1002–1017.

Pechmann, T. 1989. Incremental speech production and referential overspecification. *Linguistics*, 27(1):98–110.

Viethen, Jette and Robert Dale. 2011. GRE3D7: A corpus of distinguishing descriptions for objects in visual scenes. In *Proceedings of UCNLG+Eval-2011*, pages 12–22, Edinburgh.

This article has been cited by:

1. Ivandré Paraboni, Danielle Sampaio Monteiro, Alex Gwo Jen Lan. Personality-Dependent Referring Expression Generation 20-28. [Crossref]