RUNNING HEAD: VIOLATIONS OF GRICEAN MAXIMS

How do violations of Gricean maxims affect reading?

Kumiko Fukumura

School of Psychological Sciences and Health

University of Strathclyde, UK

Roger P.G. van Gompel

School of Social Sciences

University of Dundee, UK

Address correspondence to:

Kumiko Fukumura

School of Psychological Sciences and Health

University of Strathclyde

40 George Street

Glasgow G1 1 QE

United Kingdom

kumiko.fukumura@strath.ac.uk

Abstract

Four eye-tracking experiments examined how violations of the Gricean maxim of quantity affect reading. Experiments 1 and 2 showed that first-pass reading times for size-modified definite nouns (*the small towel*) were longer when the modifier was redundant, as the context contained one rather than two possible referents, whereas first-pass times for bare nouns (*the towel*) were unaffected by whether the context contained multiple referents that resulted in ambiguity. Experiment 3 showed that unlike redundant size modifiers, redundant color modifiers did not increase first-pass times. Experiment 4 confirmed this finding, demonstrating that the effect of redundancy was dependent on the meaning of the modifier. We propose that initial referential processing is led by the lexico-semantic representation of the referring expression rather than Gricean expectations about optimal informativeness: Redundancy of a size-modifier immediately disrupts comprehension because the processor fails to activate the referential contrast implied by the meaning of the modifier, whereas referential ambiguity has no immediate effect, as it allows the activation of at least one semantically-compatible referent.

Key words: ambiguity; redundancy; overspecification; Gricean maxims; reading

Introduction

The primary function of referring expressions is to identify the referent intended by the language user. Some of the principles that may guide this process are Grice's (1975) maxims of conversation. Grice's maxims make the fundamental assumption that language comprehension is led by a default expectation that an utterance should be optimally informative. Most notably, the *maxim of quantity* states that language users are expected to provide as much information as necessary but no more information than needed. The *maxim of manner* also includes submaxims such as "Avoid ambiguity" and "Be brief (avoid unnecessary prolixity)". Under Grice's maxims, if more than one bird has been mentioned in the preceding context, an unmodified definite noun such as *the bird* can be taken to be infelicitous, as it underspecifies which bird is being referred to, creating referential ambiguity. When the context contains only one bird, however, the use of a modifier, as in *the large bird*, is also in conflict with the Gricean maxims, as the modifier overspecifies the referent's properties. Although these pragmatic principles may make intuitive sense, it is not clear whether or how real-time comprehension processes might be affected by them. As we review below, the main concern in previous research has been how the referential context and the Gricean maxims help resolve syntactically ambiguous sentences. Subsequent off-line rating studies that examined the Gricean principles in the absence of syntactic ambiguity indicated that the conclusions from those studies may not generalize to sentences without syntactic ambiguity. The current study therefore focuses on the time-course with which the violations of the Gricean maxims influence online comprehension processes in syntactically unambiguous sentences, with the goal of uncovering the mechanisms that underlie referential processing more generally.

According to Grice (1975), violations of the maxims result in an inference or *conversational implicature*, whereby the literal meaning of the utterance is reconciled with the assumption that language producers are obliging the maxims. Hence, redundant descriptions may generate an implicature or inference about the language producer's rationale for the seemingly unnecessary

information (e.g., the information may be important later in the story). Similarly, ambiguous reference would also prompt comprehenders to seek for a reason behind the language user's communicative intent for it (e.g., perhaps she or he does not wish to disclose which one is intended). But crucially, Grice's theory fails to specify whether such inferential processes delay the initial comprehension of the referring expression.

Early research on on-line referential processing was instigated by research on parsing. Specifically, Crain and Steedman (1985) proposed an account, dubbed *referential theory*, which explains how the referential context affects parsing decisions. They argued that language interpretation is guided by *referential presuppositions:* for example, a definite noun phrase *presupposes* that the referent is uniquely identifiable in the context (cf. Russell, 1905; Neale, 1990). In contrast, a modified definite noun phrase presupposes a set of referents, one of which is distinguishable from the rest by the property denoted by the modifier (cf. Olson, 1970; Osgood, 1971). Note that these presuppositions are in line with, if not subsumed under, Gricean expectations about optimal informativeness (see Clifton & Ferreira, 1989; Steedman & Altmann, 1989): i.e., the use of a modifier for a definite noun would be redundant, unless we presuppose that the context contains another similar referent. Importantly, referential theory claims that when a sentence allows multiple syntactic analyses, language comprehenders *immediately* adopt an analysis, whereby these presuppositions are satisfied by the referential context. Some of the earliest evidence that supports such a claim comes from Altmann and Steedman (1988), who examined reading times for sentences such as (1):

(1a) The burglar blew open the safe with the new lock.

(1b) The burglar blew open the safe with the dynamite.

The prepositional *with*-phrase in these sentences can modify the preceding noun phrase (*the safe*) or be part of the verb phrase *blew open the safe*. But the meaning of the *with*-phrase in sentence (1a) is more compatible with the noun phrase modifier analysis, whereas the meaning of the *with*-

phrase in sentence (1b) is more compatible the verb phrase argument interpretation. Altmann and Steedman (1988) found that following a referential context that contained two safes, sentence (1b) was read more slowly than sentence (1a), which was taken to indicate that readers experienced difficulty when the meaning of the with-phrase was inconsistent with the referential context that supported the noun phrase modifier analysis. When the referential context contained only one safe, there was no reading time difference between the two sentences, though sentence (1b) was read faster in the one-safe context, where the verb phrase argument analysis avoided redundancy, than in the two-safe context, where the same analysis led to referential ambiguity. By adopting finer temporal measures, subsequent studies showed, in both written (e.g., Altmann, Garnham, & Dennis, 1992; Altmann, Garnham, & Henstra, 1994; Van Berkum, Brown, & Hagoort, 1999) and spoken language comprehension (e.g., Chambers, Tanenhaus, & Magnuson, 2004; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Spivey, Tanenhaus, Eberhard, & Sedivy, 2002), that these effects occur *as soon as* comprehenders encounter the relevant referring expressions. For instance, Tanenhaus et al. (1995) recorded listeners' eye fixations to objects in the visual scene when they listened to instructions to carry out an action on those objects. When instructions such as (2a) were presented with a visual context of only one apple, an empty towel in the scene received increased fixations, which were time-locked to the onset of "on the towel", suggesting that listeners initially analysed this phrase as the destination of the apple rather than as a modifier of "the apple". Crucially, when the visual display contained two apples, one on a towel and the other on a napkin, there was no increase in the fixations to the empty towel following (2a) as compared to fixations in the syntactically unambiguous instruction (2b). These findings were taken to indicate that redundancy of "on the towel" in the one-referent context led to the misanalysis of the prepositional modifier as the destination of the instructed action, whereas referential ambiguity in the two-referent context favoured the noun phrase modifier interpretation, which in turn helped the correct parsing decision.

(2a) Put the apple on the towel in the box.

(2b) Put the apple that's on the towel in the box.

The strong version of the Gricean hypothesis

Although these earlier studies were motivated by questions concerning syntactic processing, the important implications are that language comprehenders have fairly strong Gricean expectations about optimal informativeness, such that the referential context predisposes comprehenders to certain referring expressions over others, which is why the referential context exerts an immediate influence on syntactic analyses involving different referential forms. This raises the possibility that even in the absence of syntactic ambiguity, violations of Gricean expectations should disrupt comprehension immediately; upon encountering referring expressions that violate Gricean expectations because they provide too little or too much information, comprehenders will immediately experience difficulty. This assumed immediacy is an extension of Grice's (1975) original proposal, which did not specify exactly how referentially ambiguous or redundant descriptions should impair initial comprehension processes. Hence, we call it the *strong version* of the Gricean hypothesis.

However, other studies have shown that the effects of referential context on syntactic analyses are less strong (Ferreira & Clifton, 1986; Murray & Liversedge, 1994; Zagar, Pynte, & Rativeau, 1997; Spivey & Tanenhaus, 1998). Though such findings have been taken to support the view of the independence of syntactic processing from pragmatic constraints (e.g., Frazier, 1987), they also cast doubt over the strength of Gricean expectations and their impact during initial comprehension processes. Furthermore, recent off-line rating studies as well as research on anaphoric processing suggest that violations of the Gricean constraints do not always hinder comprehension. In the current study, we therefore propose and test two alternative hypotheses concerning how the different violations of the Gricean maxims might influence online comprehension processes, which we shall now discuss in turn.

Ambiguity first hypothesis

Using a similar set-up as in Tanenhaus et al. (1995), Engelhardt, Bailey, and Ferreira (2006) had participants rate the felicity of different spoken instructions in one of their experiments. Participants rated instructions that contained referential ambiguity ("Put the apple in the box" in the context of two apples) as less appropriate than instructions with a disambiguating post-nominal modifier (e.g., *Put the apple on the towel in the box* in the context of two apples), whilst they did not reliably rate instructions with a redundant modifier (e.g., *Put the apple on the towel in the box* in the context with only one apple) as less appropriate than instructions with no redundancy (*Put the apple in the box* in the context of only one apple). In a similar rating study, Davies and Katsos (2013) found that participants rated *both* ambiguous and redundant expressions as less natural than expressions that were neither ambiguous nor redundant, though participants rated ambiguous expressions as less natural than redundant expressions. Davies and Katsos argued that with ambiguous descriptions, the referent cannot be identified uniquely, whereas redundant descriptions do allow unique identification, so ambiguity is more problematic to comprehenders than redundancy. Furthermore, Arts, Maes, Noordman and Jansen (2011) found that redundancy does not always hinder object identification; in some cases, it can even *facilitate* comprehension. For instance, when the description *round button* identified the referent uniquely, *round white button* neither hindered nor facilitated referent identification, and highly redundant descriptions like *large round white button* led to shorter identification times relative to *round button*.

Whilst the extent to which these studies inform us about online comprehension processes is not entirely clear, it is possible that comprehension difficulty arises primarily when the processor fails to identify the referent uniquely in the context. If so, referential ambiguity may affect comprehension more strongly and perhaps more rapidly than redundancy. We may thus hypothesize that as soon as comprehenders encounter a referring expression, they immediately look for a referent that *uniquely* matches the meaning of the referring expression. If the context contains more than one

possible referent and therefore it is not possible to identify the referent unambiguously, comprehenders will quickly experience difficulty. In other words, comprehenders will not, at least initially, experience difficulty, as long as the referring expression identifies the referent uniquely, so redundant referring expressions will disrupt comprehension more slowly than ambiguous referring expressions. We call this the *ambiguity-first* hypothesis.

Redundancy first hypothesis

Other evidence, however, suggests that ambiguity does not immediately influence comprehension processes. Several models of anaphoric processing (Cook & O'Brien, 2014; Garrod et al., 1990; Gernsbacher, 1989; Rigalleau et al. 2004) assume that initial anaphoric processing is guided by the degree of semantic fit between the antecedent and the anaphoric expression: Upon encountering an anaphor, the antecedent that best fits the semantic specification of the anaphor is most strongly activated. If multiple referents are consistent with the description of an anaphor, they will all be simultaneously active, in keeping with models of memory retrieval (e.g., Hintzman, 1986; Ratcliff, 1978) that posit parallel activation of multiple traces for a given probe, with traces more similar to the probe being more strongly activated. The parallel activation of multiple referents may mean that the initial referent-expression mapping does not need to be unique, such that ambiguous descriptions will not immediately disrupt comprehension.

Indeed, research on anaphor resolution has shown that gender ambiguous pronouns (e.g., *she* in the context of multiple female referents) does not disrupt comprehension immediately (e.g., Garnham, Oakhill, & Cruttenden, 1992; Greene, McKoon, & Ratcliff, 1992; Rigalleau, Caplan & Baudiffier, 2004; Stewart, Holler, & Kidd, 2007, but see also Niewland, Otten, & Van Berkum, 2007). Research by Rigalleau and Caplan (2000) even suggests that under some conditions, gender ambiguous pronouns are processed faster than gender unambiguous pronouns, with one possible reason being that gender ambiguous pronouns activate all the possible referents consistent with the pronoun's gender (cf. Garrod, O'Brien, Morris, & Rayner, 1990; Gernsbacher, 1989), whereas

gender unambiguous pronouns deactivate gender-incompatible referential alternatives, and this *disengagement* process takes time. Similarly, research has shown that syntactic ambiguity does not necessarily slow down comprehension (e.g., Logačev & Vasishth, 2016; Swets, Desmet, Clifton, & Ferreira, 2008; Traxler, Pickering, & Clifton, 1998; Van Gompel, Pickering & Traxler, 2001).

Hence, contra the ambiguity first hypothesis, referential ambiguity may not delay comprehension immediately. That is, although comprehenders may promptly seek to identify a referent, difficulty may not arise *unless* they fail to identify *at least one referent* that is compatible with the description of the referring expression. Ambiguous definite nouns, such as *the bird* in the context of multiple birds, allow multiple referent-expression mappings, so the initial referential processing will proceed without any delay. Interestingly, under this account, redundant descriptions *could* have an immediate impact. Research has shown that comprehenders immediately interpret modifiers contrastively (e.g., Hanna, Tanenhaus, & Trueswell, 2003; Sedivy, 2003; Sedivy, Tanenhaus, Chambers, & Carlson, 1999). That is, upon encountering *the large bird*, comprehenders immediately consider not only the referent of the expression (the large bird), but also other referents that are implied by the modifier (smaller birds). When such an expression occurs in the context of only one bird, then comprehenders should experience immediate difficulty, as they fail to find the smaller birds implied by the modifier. We call this third hypothesis the *redundancy-first* hypothesis.

Evidence for this hypothesis comes from Engelhardt, Demiral, and Ferreira (2011), who found that redundant modifiers impair comprehension. In their study, participants were shown a visual display of two objects, before listening to an instruction containing either size- or color-modified descriptions, as in *Look at the big/red star*. Participants were slower at indicating on which side of the display (right or left) the target object was presented when size or color modifiers were redundant in the display (e.g., the two objects had different nominal categories) than when they were not (e.g., the objects had the same nominal categories). Moreover, their ERP data showed more negative ERP responses 450-570ms after the onset of redundant adjectives. Thus, the study provided

evidence that redundancy does impair comprehension. Yet these findings are in conflict with those from Arts et al. (2011), who showed that redundancy can facilitate comprehension. Also, because Engelhardt et al. did not examine the impact of referential ambiguity, we do not know how the time-course of the redundancy effect compares to that of ambiguity.

In sum, the current study examined whether and how the Gricean maxims might affect referential processing during reading. The first two experiments examined a strong version of the Gricean hypothesis against the two alternative hypotheses introduced above, which claim that initial comprehension processes are only sensitive to either referential ambiguity or redundancy. As a preview, the first two experiments showed that whereas referential redundancy led to an immediate slow down, referential ambiguity did not. We then carried out additional two experiments to examine the nature of these findings.

Experiment 1

The first experiment examined the time-course with which referential ambiguity and referential redundancy affect comprehension. We recorded participants' eye movements while they read sentence pairs such as (3). The first sentence, which we call the *context sentence*, set the referential context, introducing the relevant discourse referents. The second sentence, which was the *target sentence*, then referred to one of those entities, with its grammatical subject being realized as either a definite bare noun (*the towel,* 3a & 3b) or a definite noun modified by a size adjective (*the small towel,* 3c & 3d). In the *one-referent* context (3a & 3c), the context sentence mentioned one exemplar of each of two different nominal categories (*towel* and *robe*), so the context contained only one possible referent for the nominal category denoted by the subject in the target sentence. In the *two-referent* context (3b & 3d), the sentence mentioned two exemplars of the same nominal category (*towel*), so there were two referents that had the same nominal category as the subject in the target sentence. In the two-referent context, the use of a bare noun in the target sentence resulted in

underspecification or ambiguity (3b). In the one-referent context, a modified noun in the target sentence was overspecified, as the modifier was redundant for referent identification (3c).

(3a) There was a small towel and a large robe in the bathroom. The towel was soaking on the floor.
(3b) There was a small towel and a large towel in the bathroom. The towel was soaking on the floor.
(3c) There was a small towel and a large robe in the bathroom. The small towel was soaking on the floor.
(3d) There was a small towel and a large towel in the bathroom. The small towel was soaking on the floor.

Our main interest was in how long readers spent reading the referring expression and the following words in the target sentence as a function of the referential context set by the context sentence. To determine the time course of different effects on referential processing, we will distinguish between first-pass reading times for the referring expressions (i.e., the fixations on the referring expression before readers exit the region) reflecting very rapid effects and later eye movement measures such as total times for the expression and reading times for words following it, in keeping with previous studies (e.g., Garrod & Terras, 2000). The strong version of the Gricean hypothesis predicts that both referential ambiguity and referential redundancy should immediately disrupt comprehension. Hence, first-pass times for the bare noun in the two-referent context (3b) should be longer than in the one-referent context (3a), whereas first-pass times for the modified noun should be longer in the one-referent context (3c) than in the two-referent context (3d). However, early referential processing may not be driven by Gricean expectations. The *ambiguity-first* hypothesis assumes that referential ambiguity affects reading difficulty immediately, whereas referential redundancy does not, because initial comprehension difficulty arises primarily when the referring expression fails to identify its referent uniquely. That is, while there should be increased first-pass reading times of the bare noun in the two-referent context (3b) relative to the one-referent context (3a), first-pass reading times for the modified noun should be unaffected by the referential context; instead, the effect of the redundant modifiers may emerge in late reading time measures such as total times for the referring expression or regressive eye movements following the referring expression. Alternatively,

the *redundancy-first* hypothesis assumes that early referential processing is not constrained by referential uniqueness; instead initial processing focuses on identifying referents relevant for the meaning of the referring expression. On this account, redundant modifiers result in an immediate slowdown in reading, because the processor fails to identify the contrast set implied by the referring expression, whilst ambiguous bare nouns allow the identification of at least one referent. So while redundancy should slow down first-pass times for the modified noun, ambiguity should not influence first-pass times for the bare noun; the effect may arise in total times for the bare noun and/or in regressive eye movements from later parts of the sentence.

Method

Participants. Forty native speakers of British English (32 females), aged between 18 and 30, were recruited from the University of Strathclyde student community for course credit. Data from two participants were replaced due to too many track losses. All participants had normal or corrected-to-normal vision and reported to be non-dyslexic.

Materials. We constructed 48 experimental items such as (3), each of which consisted of two sentences presented on a single line. The first sentence served as the *context sentence* and began with *There was*, followed by a noun phrase coordination involving two singular nouns and a prepositional phrase (e.g., *in the bathroom*). The second sentence, which was the *target sentence*, referred to one of the referents in the coordinated noun phrase, and its subject was always singular and realized as a definite noun phrase either with (3c & 3d) or without (3a & 3b) the modifier used in the context sentence. In the *one-referent* context (3a & 3c), the context sentence introduced two nouns with different nominal categories (e.g., a towel and a robe), so the context contained only one possible referent for the subject of the target sentence. In the *two-referent* context (3b & 3d), the two nouns in the context sentence had the same nominal category (e.g., two towels), so there were two possible referents for the subject of the target sentence. In both contexts, the nouns were modified by size adjectives (e.g., *small*, *large*). The noun without the modifier (bare noun) was ambiguous or

underspecified in the two-referent context (3b), but not in the one-referent context (3a). The modified noun was redundant or overspecified in the one-referent context (3c), but not in the two-referent context (3d). See appendix for the list of experimental materials. In addition, we constructed 112 filler items and 8 practice items. The filler items had varied sentence structures: some items had similar constructions to the experimental items (e.g., *There were many tents put up in the field. It was because of the festival*) and others had completely different constructions (*The athletes all trained hard during the year. They were looking forward to the Olympics*). No filler items contained ambiguous or redundant descriptions.

Design. The *referring expression* (bare noun vs. modified noun) in the target sentence and the *referential context* (one referent vs. two referents) in the context sentence were orthogonally manipulated. In addition, which of the two nouns mentioned in the context sentence (first vs. second) was referred to in the target sentence was included as a within-participants and within-items counterbalancing variable. This resulted in eight conditions in total. The 48 experimental items along with the eight practice items and 112 filler items were pseudo-randomly distributed across eight lists, with six items from each condition and one version from each item, subject to the constraint that there were at least two filler items between the experimental items. Five participants were randomly assigned to each list.

Procedure. Participants' eye movements were recorded using an Eyelink 1000 eye tracker. The eye fixations were sampled at 1000Hz and the tracker had an angular resolution of .02 degree. Participants were seated at a distance of 58cm from a 19inch CRT display (1600 × 1200 pixel resolution). They were instructed to stay as still as possible, and rested on a chin and head rest to reduce head movements. Before the initial trial, at least one calibration was carried out and further calibrations were performed throughout the experiment at regular predefined intervals and when required. In each trial, participants were instructed to fixate on a dot, which appeared in the centre of the screen, until the experimenter had checked the accuracy and manually accepted the fixation.

Participants were then instructed to fixate on a cross, appearing at the left edge of the screen. A fixation on this cross triggered the presentation of the sentence. All the sentences were presented on a single line in 20-point Times New Roman font. Participants were asked to read the sentence carefully at their normal speed, and then to press a button on a game pad, which they held in their hands. In some trials (25 experimental trials, 80 filler trials), this prompted the presentation of a comprehension question (e.g., *Where was the towel?*). Participants were asked to choose one of the possible answers presented either on the right or left side of the screen using the game pad (e.g., *On the floor* vs. *On the hook*). About half the comprehension questions asked about the context sentence and the other half about the target sentence. Comprehension questions that would highlight the ambiguity of bare nouns in the experimental trials were avoided. The experiment took about 45 minutes, with participants being offered regular breaks before predefined calibrations.

Results

The average accuracy rate for the comprehension questions was 98%, indicating that participants understood the sentences. We analysed the reading times for the target sentence by dividing the sentence into the following three regions, shown by brackets in (4).

(4) [The towel/ The small towel][was soaking] [on the floor.]

The first region was the *critical region*, which comprised the subject noun phrase in the target sentence. The space before the referring expression was included in the critical region. The second region was the *post-critical region*, comprising the word(s) immediately following the critical region. This region included further words if the region was shorter than eight characters. The third region was the final region, comprising the remainder of the sentence, which also had at least eight characters.

Fixations shorter than 40ms were combined if they were within one character space of the previous or next fixation. Remaining fixations shorter than 80ms or longer than 1000ms were discarded. Before the analyses, trials with track losses or blinks that occurred during first-pass in a

particular region were excluded together with other errors (e.g., participants pressed a key to move on without reading the sentences). This led to the exclusion of 4.4% of total trials in the critical region, and the exclusion rates were comparable across conditions: 4.3% in the bare noun/one referent condition, 4.3% in the bare noun/two referent condition, 5.1% in the modified noun/one referent condition; 4.0% in the modified noun/two referents condition.

We employed four eye-movement measures to analyse each region. *First-pass time* is the sum of fixations in a region before readers exited the region for the first time, either to the left or right, provided that they had not fixated any subsequent region. Because this measure only includes fixations that occur when participants see the region for the first time, it provides an early measure of reading difficulty. *First-pass regressions* is the percentage of trials on which readers make regressive eye movements following a first-pass fixation in the region. *Regression-path time* is the sum of fixations in the region before readers go past the region, again provided that no following region has been fixated before. *Total time* is the sum of all fixations within the region, which includes fixations during re-reading of the region. Reading times with z-scores exceeding 4 were removed, resulting in the removal of no more than 1.0% of data in any measure for a region. For all reading-time measures, trials on which participants did not fixate the region were excluded from the analyses (i.e., they were not treated as 0ms). The means of the eye-movement measures are reported in Table 1.

Throughout this article, we carried out linear mixed effects analyses (Baayen, Davidson, & Bates, 2008; Jaeger, 2008) on first-pass time, regression path duration and total time using the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015) of the statistical software R (version 3.3.1: R Development Core Team, 2016). The *p*-values were based on Satterthwaite's approximations, calculated by the *ImerTest* package. First-pass regressions were analysed using logit mixed effect models. The analyses included *referential context* (one- vs. two-referent context) and *referring expression* (bare nouns vs. size-modified nouns) as mean-centred, fixed effects. We always included by-participants and by-items random intercepts and their slopes for all the relevant fixed effects

(Barr, Levy, Scheepers, & Tily, 2013), by enforcing zero correlations between random effects in order to avoid overparameterization or false convergence (Bates, Kliegl, Vasishth, & Baayen, 2015).

Our primary goal was to determine whether referential context affected the reading of both unmodified bare nouns and size-modified nouns, and whether the referential context differentially influenced each referring expression. Hence, we focus our discussion on the simple effects of referential context on each referring expression and the referential context × referring expression interaction. The main effects of referential context and referring expression are reported in Table 2, along with the simple effects of context and the referential context × referring expression interaction.

Table 1. Means of eye movement measures by region in Experiment 1

		Region		
Expression	Context	Critical	Post-critical	Final
First-pass time (ms)				
Bare noun	One referent	269 (5)	357 (8)	380 (11)
Bare noun	Two referents	258 (5)	364 (8)	392 (12)
Size-modified noun	One referent	430 (9)	327 (7)	389 (11)
Size-modified noun	Two referents	384 (7)	332 (8)	394 (11)
First-pass regressions (%)				
Bare noun	One referent	3.9 (1.0)	14.5 (1.7)	83.6 (1.8)
Bare noun	Two referents	2.6 (0.8)	14.9 (1.7)	81.8 (1.8)
Size-modified noun	One referent	4.2 (0.9)	10.1 (1.4)	85.7 (1.7)
Size-modified noun	Two referents	3.0 (0.8)	11.0 (1.5)	81.7 (1.9)
Regression-path time (ms)				
Bare noun	One referent	286 (7)	433 (13)	1362 (53)
Bare noun	Two referents	266 (5)	435 (12)	1293 (48)
Size-modified noun	One referent	460 (11)	381 (11)	1226 (44)
Size-modified noun	Two referents	404 (9)	396 (12)	1086 (38)
Total time (ms)				
Bare noun	One referent	415 (11)	554 (15)	512 (15)
Bare noun	Two referents	424 (13)	595 (16)	529 (15)
Size-modified noun	One referent	610 (15)	474 (13)	494 (15)
Size-modified noun	Two referents	576 (16)	486 (13)	511 (15)

Note: standard errors in brackets.

Table 2. Analyses of eye movement measures by region in Experiment 1

		Critical region				*Post-critical region*				*Final region*			
		β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
First-pass time	Expression	73.68	5.50	13.40	<.001*	-15.60	4.70	-3.32	.002*	4.06	4.69	0.87	.386
	Context	-15.19	2.92	-5.20	<.001*	2.95	3.45	0.85	.393	5.17	4.85	1.07	.292
	Expression x Context	-8.13	3.08	-2.64	.012*	0.17	3.51	0.05	.961	-1.68	4.69	-0.36	.720
	Context in Bare noun	-6.14	3.22	-1.91	.063	2.93	5.46	0.54	.595	7.04	6.81	1.04	.301
	Context in Size-modified noun	-22.74	4.94	-4.60	<.001*	3.36	4.74	0.71	.479	3.34	6.83	0.49	.627
		β	*SE*	*z*	*p*	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*
First-pass regressions	Expression	0.09	0.18	0.49	.627	-0.22	0.09	-2.45	.014*	0.06	0.07	0.87	.385
	Context	-0.21	0.16	-1.33	.185	0.03	0.07	0.40	.688	-0.14	0.08	-1.85	.064
	Expression x Context	0.03	0.17	0.17	.862	0.01	0.08	0.13	.894	-0.07	0.07	-0.96	.337
	Context in Bare noun	-0.23	0.24	-0.95	.343	0.01	0.10	0.10	.918	-0.07	0.10	-0.73	.469
	Context in Size-modified noun	-0.19	0.24	-0.77	.440	0.04	0.12	0.36	.721	-0.20	0.11	-1.89	.059
		β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Regression-path duration	Expression	80.27	6.61	12.15	<.001*	-23.90	6.77	-3.53	.001*	-86.44	19.28	-4.48	<.001*
	Context	-19.56	4.06	-4.82	<.001*	3.97	5.48	0.72	.474	-56.55	25.87	-2.19	.035*
	Expression x Context	-8.25	4.12	-2.00	.051	3.53	5.71	0.62	.542	-15.01	22.98	-0.65	.518
	Context in Bare noun	-10.86	3.99	-2.72	.010*	0.49	8.41	0.06	.954	-42.19	38.68	-1.09	.282
	Context in Size-modified noun	-27.19	6.62	-4.11	<.001*	7.78	8.59	0.91	.371	-71.75	29.97	-2.39	.022*
		β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Total time	Expression	90.74	8.22	11.04	<.001*	-48.42	6.84	-7.08	<.001*	-8.46	6.55	-1.29	.208
	Context	-7.66	7.03	-1.09	.285	11.92	8.04	1.48	.147	9.78	6.13	1.60	.119
	Expression x Context	-10.67	6.08	-1.76	.086	-6.49	6.52	-1.00	.330	-0.46	6.91	-0.07	.948
	Context in Bare noun	2.49	9.81	0.25	.801	18.52	11.28	1.64	.111	10.36	8.58	1.21	.234
	Context in Size-modified noun	-17.84	9.30	-1.92	.063	5.41	9.58	0.57	.577	8.78	9.81	0.90	.377

Critical region (The towel/the small towel). In this region, referential context × expression interactions emerged in first-pass times and regression-path times. First, first-pass times for modified nouns were read 46ms faster in the two-referent condition than in the one-referent condition, whilst no significant effect was found in first-pass times for bare nouns. Second, regression-path times were shorter in the two-referent condition than in the one-referent condition in both the bare noun (20ms) and in the modified noun condition (56ms), but this effect of referential context tended to be larger with modified nouns than with bare nouns. Third, total times for modified nouns were marginally shorter in the two-referent condition than in the one-referent condition, while no significant context effect was found in total times for bare nouns. No significant context effects were found in first-pass regressions for either referring expression.

Extended critical region (The towel was). Due to the additional adjective, modified noun phrases are longer than bare noun phrases. One possibility was that the critical region for bare nouns was too short to obtain a referential context effect; because the bare noun region was short, participants often fixated only once in this region (79%, as opposed to 34% for modified nouns), so first-pass times for bare nouns were often equivalent to first-fixation durations. As a result, the first-pass measure, mostly comprising only the very first fixations, might have been too early to detect a referential context effect. We thus extended the bare noun region by one word (*the towel was*), lowering the rate of single fixations (as opposed to multiple fixations) to 45%, in order to capture any delayed effects from the bare noun. No effect of context was observed on first-pass times for this extended region ($\beta = -1.45$, $SE = 5.01$, $t = -0.29$, $p = .773$). Furthermore, analyses of the extended region revealed marginally longer regression path times following the one-referent (407ms) than the two-referent (386ms) context ($\beta = -10.89$, $SE = 6.11$, $t = -1.78$, $p = .075$), and no effects on regressions ($\beta = -0.27$, $SE = 0.19$, $z = -1.46$, $p = .145$), and total times ($\beta = 18.14$, $SE = 12.54$, $t = 1.45$, $p = .158$), suggesting that ambiguity of bare nouns did not cause processing difficulty.

Post critical region (was soaking). No significant context × referring expression interaction was observed in any measure and there were no simple effects of context on either bare or modified nouns.

Final region (on the floor). Regression-path times were 140ms shorter when the modified nouns occurred in the two-referent condition than in the one-referent condition, whereas the referential context did not reliably influence regression-path times following bare nouns, though the context × expression interaction was not significant. No other context effects were significant in this region.

Discussion

The first-pass times for the critical region showed a clear effect of referential context on modified nouns; readers read modified nouns more slowly in the one-referent context than in the two-referent context, indicating that redundancy immediately slowed down reading. Evidence for processing difficulty with modified nouns in the two-referent context also emerged in regression-path times in the critical and final regions. In contrast, referential ambiguity did not affect reading of the bare noun; if anything, the first-pass times for bare nouns were marginally *shorter* in the two-referent context, where the use of a bare noun was referentially ambiguous, than in the one-referent condition, where the bare noun was unambiguous. The regression-path times for the bare nouns were also *shorter* in the two-referent condition than in the one-referent condition. Hence, the results show no evidence that ambiguity disrupts comprehension, in contrast to redundancy.

One possibility, which was raised during the review process, was that there might have been a default preference to interpret the subject noun in the target sentence as co-referential with one of the entities introduced in the context sentence. For instance, readers might have preferentially interpreted the subject of the target sentence as being co-referential with the first-mentioned antecedent in the context sentence (*the small towel* in 3) rather than the second-mentioned antecedent (*the large towel/robe* in 3). If so, the bare subject noun in the target sentence was essentially unambiguous. Because we counterbalanced the order of the two nouns in the context sentence, we could examine

whether the position of the referent in the context sentence (first or second) had an effect on the bare nouns in the one-referent (unambiguous) condition. No such effect was found, $\beta = 1.05$, $SE = 5.35$, $t = 0.20$, $p = .846$. Hence, it is very unlikely that order of mention biased the interpretation of ambiguous bare nouns.

Another concern was that in the one-referent condition, the head noun of the target description was mentioned only once in the context sentence, whereas it was mentioned twice in the two-referent condition. This was done in line with previous reading studies such as Altmann and Steedman (1988), where the referential candidates were individually labelled and distinguished in the preceding sentence. But the repeated mention of the head noun in the two-referent context may have resulted in stronger lexical priming, which might have enhanced the redundancy effect of the modified noun and offset the cost of referential ambiguity for bare nouns. Although this would mean that the cost of referential ambiguity was not sufficiently strong to overrule the effect of lexical priming, we need to know if different results emerge when lexical priming is ruled out. The aim of Experiment 2 was to address this concern.

Experiment 2

Unlike in the first experiment, the context sentence in the second experiment involved a single noun phrase, as shown in (5). In the *one-referent* context (5a & 5c), the noun phrase was preceded by the numeral *one*, whereas in the *two-referent* context (5b & 5d), the numeral *two* preceded the noun phrase. Hence, the critical noun (*towel*) was mentioned once in both the one-referent and two-referent context. In both contexts, the target sentence started with either an unmodified definite noun phrase (5a & 5b) or a size-modified definite noun phrase (5c & 5d). If readers interpret these referring expressions in the target sentence in accord with the maxim of quantity, the bare noun should be read more slowly following the two-referent context (5b) (relative to the one-referent context, 5a), where the bare noun underspecifies which of the two referents in the context sentence is referred to. In contrast, the modified noun should be read more slowly in the one-referent context

(5c) (relative to the two-referent context, 5d), where the modifier overspecifies the referent. The strong version of the Gricean account predicts that both referential ambiguity and referential redundancy should increase first-pass reading time for bare nouns and modified nouns respectively. In contrast, the ambiguity-first hypothesis predicts that referential ambiguity should affect comprehension more quickly than referential redundancy; hence although referential ambiguity should slow down first-pass reading times for bare nouns, redundancy should not affect first-pass times for modified nouns; slowdown due to redundancy may be reflected in later measures such as regression-path times or total times for the referring expressions and reading times in later regions. In contrast, the redundancy-first hypothesis predicts that the opposite should be the case.

(5a) There was one towel on the floor. The towel was soaking on the floor.

(5b) There were two towels on the floor. The towel was soaking on the floor.

(5c) There was one towel on the floor. The small towel was soaking on the floor.

(5d) There were two towels on the floor. The small towel was soaking on the floor.

Method

Participants. Forty new participants (30 females), recruited from the same population as in Experiment 1, took part. The data from one participant was replaced with a new participant due to too many track losses and another participant was replaced due to too many errors (52%) on the comprehension questions.

Materials. The materials were the same as in Experiment 1, except that the context sentence explicitly mentioned the number of referents using numerals, as in (5). See appendix.

Design. There were four conditions in total: *referring expression* (bare noun vs. modified noun) × *referential context* (one-referent vs. two-referent). This resulted in four lists. In each list, the 48 experimental items along with 8 practice items and 112 filler items were randomly distributed. Each list consisted of 12 items from each condition and one version from each item. Ten participants were randomly assigned to each list.

Procedure. This was the same as in Experiment 1.

Results

Participants answered the comprehension questions very accurately (99%). Fixation data were treated as before, which led to the exclusion of 4.8% of total trials in the critical region (4.8% in the bare noun/one referent condition, 5.8% in the bare noun/two referents condition, 4.0% in the modified noun/one referent condition and 4.9% in the modified noun/two referent condition). The data were analysed in the same way as in Experiment 1. Table 3 reports the means of the reading measures for the target sentence, and Table 4 the results of the analyses.

Table 3. Means of eye movement measures by region in Experiment 2

		Region		
Expression	Context	Critical	Post-critical	Final
First-pass time (ms)				
Bare noun	One referent	273 (5)	351 (8)	420 (11)
Bare noun	Two referents	281 (6)	362 (8)	406 (12)
Size-modified noun	One referent	473 (10)	357 (8)	429 (13)
Size-modified noun	Two referents	446 (9)	347 (7)	414 (12)
First-pass regressions (%)				
Bare noun	One referent	0.7 (0.4)	10.0 (1.4)	55.0 (2.4)
Bare noun	Two referents	1.1 (0.5)	17.8 (1.8)	64.4 (2.3)
Size-modified noun	One referent	3.0 (0.8)	17.1 (1.7)	67.0 (2.2)
Size-modified noun	Two referents	2.0 (0.7)	10.3 (1.4)	61.2 (2.3)
Regression-path time (ms)				
Bare noun	One referent	272 (5)	391 (10)	796 (27)
Bare noun	Two referents	286 (6)	454 (13)	984 (34)
Size-modified noun	One referent	500 (12)	465 (14)	1036 (39)
Size-modified noun	Two referents	456 (9)	390 (9)	869 (30)
Total time (ms)				
Bare noun	One referent	333 (8)	474 (12)	497 (13)
Bare noun	Two referents	425 (13)	531 (14)	524 (15)
Size-modified noun	One referent	698 (17)	485 (12)	537 (16)
Size-modified noun	Two referents	583 (13)	459 (12)	499 (14)

Note: standard errors in brackets.

Table 4. Analyses of eye movement measures by region in Experiment 2.

		Critical region				*Post-critical region*				*Final region*			
		β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
First-pass time	Expression	95.13	6.88	13.83	<.001*	-2.67	3.60	-0.74	.462	4.68	6.32	0.74	.463
	Context	-5.60	3.31	-1.69	.091	0.26	3.80	0.07	.947	-5.23	5.95	-0.88	.384
	Expression x Context	-8.84	3.34	-2.65	.011*	-5.52	3.66	-1.51	.139	0.36	5.34	0.07	.946
	Context in Bare noun	3.92	4.00	0.98	.337	5.91	5.74	1.03	.312	-5.38	7.94	-0.68	.504
	Context in Size-modified noun	-14.10	5.45	-2.59	.010*	-5.19	4.90	-1.06	.289	-4.84	7.56	-0.64	.523
		β	*SE*	*z*	*p*	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*
First-pass regressions	Expression	-0.06	1.14	-0.06	.956	< 0.01	0.08	-0.03	.978	0.12	0.06	1.97	.049*
	Context	-0.56	1.10	-0.51	.612	0.03	0.07	0.39	.700	0.05	0.06	0.81	.416
	Expression x Context	0.15	1.13	0.13	.893	-0.36	0.07	-4.82	<.001*	-0.26	0.08	-3.14	.002*
	Context in Bare noun	-1.06	2.42	-0.44	.662	0.39	0.11	3.60	<.001*	0.31	0.11	2.89	.004*
	Context in Size-modified noun	-0.25	0.35	-0.73	.465	-0.34	0.11	-3.01	.003*	-0.22	0.11	-2.12	.034*
		β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Regression-path time	Expression	102.73	6.93	14.83	<.001*	1.19	7.05	0.17	.866	29.42	14.31	2.06	.048*
	Context	-8.53	3.76	-2.27	.023*	-2.38	6.05	-0.39	.697	4.67	12.99	0.36	.720
	Expression x Context	-14.56	3.76	-3.87	<.001*	-34.88	8.09	-4.31	<.001*	-91.75	16.53	-5.55	<.001*
	Context in Bare noun	6.80	3.98	1.71	.100	32.54	9.35	3.48	.002*	99.26	20.18	4.92	<.001*
	Context in Size-modified noun	-22.49	6.37	-3.53	<.001*	-36.79	10.03	-3.67	.001*	-86.14	21.26	-4.05	<.001*
		β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Total time	Expression	133.45	8.28	16.11	<.001*	-14.74	6.03	-2.45	.018*	5.03	5.86	0.86	.392
	Context	-6.03	5.81	-1.04	.305	7.62	6.12	1.25	.224	-1.87	5.61	-0.33	.739
	Expression x Context	-51.59	8.19	-6.30	<.001*	-22.26	5.66	-3.93	.001*	-14.93	6.21	-2.40	.020*
	Context in Bare noun	46.10	8.67	5.32	<.001*	30.11	8.73	3.45	.002*	13.35	8.18	1.63	.111
	Context in Size-modified noun	-57.57	9.46	-6.09	<.001*	-14.56	7.89	-1.85	.078	-17.49	8.57	-2.04	.043*

Critical region (The towel/the small towel). We found significant context × expression interactions in first-pass times, regression-path times and total times. First-pass times for the modified nouns were 27ms slower in the one-referent condition than in the two-referent condition. In contrast, first-pass times for the bare nouns were unaffected by the referential context. Likewise, regression-path times for modified nouns were 44ms longer in the one-referent condition than in the two-referent condition, whereas no reliable effect of context was found for bare nouns. Finally, whereas total times for the modified nouns were 115ms shorter in the two-referent condition than in the one-referent condition, they were 92ms longer in the two-referent condition than in the one-referent condition for bare nouns. No significant effects were found in first-pass regressions.

Extended critical region (The towel was). As in Experiment 1, we looked for a referential context effect by extending the bare noun region by one word (*the towel was*). First-pass times for bare nouns did not differ significantly between the two-referent (387ms) and one referent (382ms) conditions. First-pass regressions, $\beta = 0.30$, $SE = 0.51$, $z = 0.60$, $p = .549$, and regression-path times, $\beta = 8.55$, $SE = 5.29$, $t = 1.62$, $p = .113$, in the two-referent context (2.1%, 402ms, respectively) did not differ from those in the one referent context (1.1%, 386ms). Total times for the extended bare noun region were significantly longer in the two-referent condition (615ms) than in the one-referent condition (492ms), $\beta = 63.10$, $SE = 12.39$, $t = 5.09$, $p < .001$. Hence, referential ambiguity had a significant effect on total times only.

Post critical region (was always). No significant effects were found in first-pass times. However, significant cross-over interactions between context and expression were found in first-pass regressions, regression-path times and total times. First, in the bare noun conditions, participants were 7.8% more likely to regress in the two-referent condition than in the one-referent condition, whereas in the modified noun conditions, participants were 6.8% less likely to regress in the two-referent condition than in the one-referent condition. Second, regression-path times were 63ms longer in the two-referent condition than in the one-referent condition for bare nouns, but they were

75ms shorter in the two-referent condition than in the one-referent condition following modified nouns. And finally, following bare nouns, total times were 57ms longer in the two-referent condition than in the one-referent condition. Following modified expressions, total times were marginally (26ms) shorter in the two-referent condition than in the one-referent condition.

Final region (most popular). Again, no effects were found in first-pass times. First-pass regressions, regression-path times and total times showed significant context × expression interactions, however. First, in the bare noun conditions, there were 9.4% more regressions in the two-referent than in the one-referent condition, whereas in the modified noun conditions, there were 5.8% fewer regressions in the two-referent than in the one-referent condition. Second, regression-path times following bare nouns were 188ms longer in the two-referent condition than in the one-referent condition, whilst following modified nouns, regression-path times were 167ms shorter in the two-referent condition than in the one-referent condition. Finally, total times following modified nouns were 38ms faster in the two-referent than in the one-referent condition. In contrast, total times following bare nouns were not reliably affected by the referential context.

Discussion

As in Experiment 1, redundancy of the modifier immediately slowed down the reading of the modified nouns: first-pass times for the modified noun region were significantly longer in the one-referent than in the two-referent condition. Regression-path times and total times for the modifier noun region were also longer in the one-referent context. The processing difficulty due to redundancy persisted in later regions, with more first-pass regressions and longer regression-path times for the post-critical region and the final region, as well as longer total reading times in the final region. Unlike in the first experiment, the ambiguity of the bare nouns also had an effect, but the effects arose in late measures of processing difficulty: Total times for the bare noun region were longer in the two-referent condition than in the one-referent condition, and first-pass regressions, regression-path times and total times in the post-critical and final regions were also affected by

referential ambiguity. Thus, realizing that the referring expression they had just read went against their pragmatic expectations, participants launched regressive eye movements from the subsequent region, possibly to make inferences about reasons behind the pragmatic violation of the referring expression, consistent with Grice (1975). Crucially, such inferences took place relatively late, as first-pass times for the bare noun region were unaffected by the referential context.

Hence, although redundant modifiers and ambiguous bare nouns both disrupted comprehension, only redundancy affected the initial reading times for the referring expressions, providing support for the redundancy-first hypothesis. On this account, referential ambiguity will not affect comprehension very quickly, because ambiguous referring expressions allow the identification of at least one referent compatible with the description of the referring expression, whereas redundant modifiers can immediately disrupt comprehension if the description fails to identify a referential alternative implied by the modifier. As discussed earlier, off-line rating studies have shown that language users consider ambiguous descriptions to be more problematic than redundant descriptions (David & Katsos, 2013) or do not even consider redundancy to be infelicitous at all (Engelhardt et al., 2006). One possibility is that these off-line ratings primarily reflect language users' concerns about communicative consequences of particular referential descriptions rather than how they process such descriptions in real time. Our results also contrast with some research on parsing (e.g., Tanenhaus et al., 1995; Van Berkum et al., 1999), which demonstrated an immediate preference for syntactic analyses that avoid both referential ambiguity and redundancy, though the results are consistent with evidence that the referential context does not immediately influence sentence processing (Ferreira & Clifton, 1986; Murray & Liversedge, 1994; Zagar et al., 1997; Spivey & Tanenhaus, 1998).

Unlike in Experiment 1, the two referential candidates in the two-referent context were introduced by a single noun phrase (e.g., *the two towels*), which avoided the repetition of the nouns in the two-referent context that we had in Experiment 1. Some theories of reference, however,

assume that for any discourse entity to be represented by language users, it must be introduced by a separate noun phrase (Prince, 1981). Under such an account, plural noun phrases with a numeral, such as *the two towels*, used in the current study, may be represented as a set of two related elements (Van Deemter, 2002). If so, it should require a *bridging inference* to refer back to one of the elements (Poesio, 2003). Also, there may have been other reasons why readers had to perform more inferences in the two-referent than the one-referent context. For example, in the two-referent context example in (5d), readers may infer at the modified noun that because the towel mentioned in the target sentence was small, the other towel was large. However, these accounts predict *longer* reading times for modified noun phrases in the two-referent than one-referent condition. Yet our results showed that reading times were *shorter* in the two-referent condition. Thus, although it is possible that readers made an additional inference in the two-referent condition, this does not explain our results.

Experiment 3

The first two experiments showed that redundancy of modifiers immediately slowed down reading, whereas ambiguity of bare nouns did not. Here we explore the mechanisms that give rise to these findings. Grice (1975) noted that whilst redundancy may simply be a "waste of time", it may also trigger inferences because comprehenders might assume that the additional information is meaningful, though Grice was unclear if such inferences would delay the initial processing of redundant descriptions. In referential theory (Crain & Steedman, 1985), the redundancy effect is attributed to a violation of the presupposition that modifiers are used to discriminate between multiple referential candidates. Crucially, referential theory assumes that referential processing proceeds incrementally, so presuppositions associated with referring expressions are taken into account on a *word-by-word* basis, which accounts for the early redundancy effects we observed in Experiments 1 and 2.

However, neither theory explains why processing difficulty with ambiguous bare nouns, which violate the Gricean maxims as well as the presupposed uniqueness constraint of definite noun

phrases, only occurred in late measures. Although this might be because readers suspended reference resolution, anticipating disambiguating information later in the sentence, such an interpretation seems at odds with the assumption of incremental referential processing assumed by referential theory. The redundancy-first-hypothesis provides an explanation for this. Under this hypothesis, referential processing slows down when the processor fails to identify not only potential target referents but also referential alternative(s) implied by the referring expression. Referential ambiguity did not immediately slow down reading, because it allowed the identification of at least one referent compatible with the description of the referring expression. By contrast, the redundancy of modifiers immediately slowed down reading, not because it violated the Gricean maxim of quantity, but rather because redundant modifiers failed to identify the relevant discourse contrast implied by the meaning of the modifier. The goal of Experiment 3 was to test this redundancy-first hypothesis by investigating the effect of redundant color modifiers.

In Experiments 1 and 2, we used size modifiers. Many researchers have argued that size adjectives are different from color adjectives. Bierwisch (1987), for instance, noted that size modifiers belong to scalar or dimensional adjectives, the representation of which necessarily involves comparison with other category exemplars; e.g., for a house to be large (as in *large house*), there must be houses that are small. Other adjectives, such as color adjectives, are only "partially gradable"; whilst the intensity of a color can vary, the color of an entity is not dependent on its comparison with other category exemplars; for houses to be red, they must possess a certain color value, but there need not be houses that are blue. This may be related to the fact that size adjectives have antonyms that specify the same dimensions in the opposite direction (e.g., large vs. small, long vs. short), whereas most color adjectives do not have obvious antonyms. Moreover, Bierwisch argued that whilst the dimension of size depends upon the meaning of the modified noun (e.g., small cup vs. small plane), the intensity of color is normally independent from the noun it modifies (e.g., red cup vs. red plane). Others have also made similar points: some properties, such as color, are more

absolute or *definite* (Martin, 1969a, 1969b) or *intrinsic* (Danks & Glucksberg, 1971) than other properties, such as size. At an empirical level, many studies have found that language users are more likely to overspecify color than size (e.g., Belke & Meyer, 2002; Pechmann, 1989; Schriefers & Pechmann, 1988). One possible reason for this is that it takes less time for language users to determine absolute properties such as color and object category than relative properties such as size (Belke & Meyer, 2002), so color is more likely to be included as part of the referent's initial conceptual representation, before language producers check its relevance for discrimination (Pechmann, 1989; Schriefers & Pechmann, 1988). Thus, although color adjectives *can* be used contrastively, they are often included as an attributive property independent of the presence of other category exemplars.

Hence, as an extension of the redundancy-first hypothesis, we propose a *meaning-based early redundancy hypothesis*. Under this account, initial referential processing is primarily led by the lexico-semantic representation of the referring expression. When encountering a referring expression, readers first activate the relevant referential candidates on the basis of the meaning of the referring expressions. Crucially, such processes occur independently from Gricean pragmatic constraints. Hence, the redundant use of a modifier can impair comprehension immediately, not because it violates the Gricean maxim of quantity, but rather because the referential context lacks the discourse entities relevant for the meaning of the modifier. The interpretation of a size modifier requires comparison between referents contrasting in size. So when encountering a size adjective, the processor automatically attempts to activate a size contrast between the target referent and alternative referents with the same nominal category that contrast in size. If the context does not provide a referential contrast, this should rapidly impair semantic processing and increase reading times of the referential description. In contrast, the color of an entity can be interpreted independently from that of other entities, so the semantic processing of color modifiers does not automatically activate a

referential contrast; hence the initial comprehension of color-modified descriptions should be unaffected by whether the context contains a referential alternative.

Several studies have shown that comprehenders interpret modifiers contrastively in the referential context (e.g., Hanna, Tanenhaus, & Trueswell, 2003; Sedivy, 2003; Sedivy et al., 1999). Sedivy et al. (1999), for instance, showed that when carrying out an action following instructions such as "Pick up the tall glass", participants were faster identifying the target referent (tall glass) when the context contained a contrasting category exemplar (small glass) than when there was no contrasting object, indicating that the presence of the other category exemplar facilitated referent identification. Interestingly, Sedivy (2003) suggested that the presence of a color-contrasted competitor did not facilitate referent identification for color-modified descriptions in an unpublished experiment. Although such a finding is inconsistent with other studies that showed that comprehenders contrastively interpret non-gradable properties (Hanna et al., 2003) and with evidence that redundant color modifiers are hard to process (Engelhardt et al., 2011), it lends provisional support to the meaning-based early redundancy hypothesis.

Thus, our third experiment examined the impact of redundant color modifiers, by contrasting the effect of referential context for unmodified and color-modified nouns, as in (6). If the early redundancy effect we observed in previous experiments was primarily led by Gricean expectations about optimal informativeness, the effect should emerge in first-pass times for the color-modified nouns; they should be slower in the one-referent context (6c) than in the two-referent context (6d). In contrast, if the early redundancy effect is specific to the meaning of size modifiers and violations of the Gricean maxims only influence later processes, then first-pass times for color modifiers should not be affected by their redundancy, though the referential context may have an effect on later eye movement measures or regions, similar to the ambiguity effect on bare nouns.

(6a) There was one towel on the floor. The towel was soaking on the floor.

(6b) There were two towels on the floor. The towel was soaking on the floor.

(6c) There was one towel on the floor. The white towel was soaking on the floor.

(6d) There were two towels on the floor. The white towel was soaking on the floor.

Method

Participants. Forty new participants (29 females) were recruited from the University of Dundee student community. As before, they were all native speakers of British English, who were aged between 18 and 30 and non-dyslexic. Four additional participants were tested but their data were discarded due to poor calibrations.

Materials and design. We had 48 experimental items and the context sentence set the referential context as in Experiment 2. However, unlike in Experiment 2, we always used color modifiers in the modified noun condition, as shown in (6). Some nominal categories used in Experiment 2 (N = 15) had to be replaced with new ones, because color could not be used to discriminate different exemplars, as the nominal categories did not have obvious color variation (e.g., hotdog, puddle, banana) (see Appendix). Hence, *referential context* (one-referent vs. two-referent) and *referring expression* (unmodified noun vs. color-modified noun) were orthogonally manipulated, resulting in four conditions. The conditions were distributed across 4 lists, each containing 48 experimental items intermixed with 112 filler items. Each list consisted of 12 items from each condition and one version from each item and ten participants were randomly assigned to each.

Procedure. This was the same as before, except that we used a 21inch monitor (1600 × 1200 pixel resolution) and participants were seated 68 cm away from it.

Results.

Before analysis, 6.9% of trials were excluded due to track losses, blinks and other technical errors (6.3% in the bare noun /one-referent condition; 6.7% in the bare noun /two-referent condition, 7.3% in the color modifier/one-referent condition and 6.9% in the color modifier/two-referent condition). Table 5 reports the means by condition. We analysed the reading times for the target sentences as a function of referring expression and referential context. Table 6 presents the results.

Critical region (*the towel/white towel*). We found a significant context × expression interaction in total times. Total times for bare nouns were 92ms slower in the two-referent context than in the one-referent context, whereas total times for colour-modified nouns were not reliably affected by the referential context. A significant context × expression interaction in regression-path times also indicated that the referential context differentially influenced regression-path times for bare nouns and colour-modified nouns, though the effect of the referential context on regression-path times for neither bare nouns nor colour-modified nouns was significant. First-pass times and first-pass regressions for both referring expressions were unaffected by the context.

Table 5. Means of eye movement measures by region in Experiment 3

		Region		
Expression	Context	Critical	Post-critical	Final
First-pass time (ms)				
Bare noun	One referent	301 (7)	369 (10)	428 (14)
Bare noun	Two referents	304 (7)	375 (9)	410 (13)
Color-modified noun	One referent	482 (10)	362 (8)	408 (12)
Color-modified noun	Two referents	479 (10)	373 (9)	400 (12)
First-pass regressions (%)				
Bare noun	One referent	2.6 (0.8)	18.1 (1.8)	59.6 (2.5)
Bare noun	Two referents	3.3 (0.9)	24.0 (2.0)	70.1 (2.3)
Color-modified noun	One referent	5.6 (1.1)	15.8 (1.7)	64.0 (2.4)
Color-modified noun	Two referents	4.3 (1.0)	11.9 (1.5)	61.6 (2.5)
Regression-path time (ms)				
Bare noun	One referent	308 (7)	464 (13)	988 (43)
Bare noun	Two referents	325 (9)	539 (18)	1179 (48)
Color-modified noun	One referent	530 (13)	457 (14)	1035 (44)
Color-modified noun	Two referents	503 (11)	431 (11)	947 (39)
Total time (ms)				
Bare noun	One referent	388 (11)	567 (15)	565 (19)
Bare noun	Two referents	480 (14)	655 (17)	566 (19)
Color-modified noun	One referent	677 (16)	548 (13)	522 (16)
Color-modified noun	Two referents	639 (15)	509 (12)	516 (16)

Note: standard errors in brackets.

Table 6. Analyses of eye movement measures by region in Experiment 3

		Critical region				*Post-critical region*				*Final region*			
		β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
First-pass time	Expression	91.97	6.28	14.64	<.001*	-3.22	4.44	-0.73	.475	-6.54	7.36	-0.89	.381
	Context	1.69	4.61	0.37	.717	4.79	3.94	1.22	.224	-5.78	6.27	-0.92	.365
	Expression x Context	-1.69	3.87	-0.44	.665	0.89	5.68	0.16	.876	2.33	6.56	0.36	.725
	Context in Bare noun	2.22	4.48	0.50	.625	3.93	6.89	0.57	.571	-7.96	10.14	-0.79	.430
	Context in Colour-modified noun	0.26	6.85	0.04	.970	6.04	5.72	1.06	.296	-2.95	7.87	-0.38	.710
		β	*SE*	*z*	*p*	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*
First-pass regressions	Expression	0.28	0.15	1.93	.054	-0.32	0.10	-3.11	.002*	-0.07	0.08	-0.91	.360
	Context	-0.02	0.15	-0.11	.911	0.02	0.08	0.23	.818	0.11	0.06	1.87	.061
	Expression x Context	-0.13	0.15	-0.88	.382	-0.21	0.08	-2.72	.007*	-0.18	0.06	-2.94	.003*
	Context in Bare noun	0.13	0.25	0.53	.594	0.23	0.09	2.49	.013*	0.29	0.09	3.31	.001*
	Context in Colour-modified noun	-0.14	0.18	-0.78	.435	-0.22	0.15	-1.42	.156	-0.07	0.09	-0.86	.388
		β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Regression-path time	Expression	102.90	8.31	12.39	<.001*	-29.24	8.98	-3.26	.003*	-45.68	18.94	-2.41	.024*
	Context	-1.44	5.50	-0.26	.795	11.83	6.37	1.86	.064	24.92	24.02	1.04	.307
	Expression x Context	-10.46	5.03	-2.08	.045*	-25.75	8.18	-3.15	.004*	-67.97	16.62	-4.09	<.001*
	Context in Bare noun	9.04	6.01	1.50	.146	37.24	11.46	3.25	.002*	97.73	31.18	3.13	.003*
	Context in Colour-modified noun	-10.74	8.54	-1.26	.217	-13.99	8.07	-1.73	.091	-42.23	27.91	-1.51	.141
		β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Total time	Expression	118.02	9.31	12.68	<.001*	-41.80	8.57	-4.88	<.001*	-21.35	9.93	-2.15	.039*
	Context	12.71	6.26	2.03	.055	12.52	7.02	1.78	.081	1.21	8.60	0.14	.889
	Expression x Context	-30.43	7.10	-4.28	<.001*	-30.54	6.03	-5.07	<.001*	-2.35	7.99	-0.30	.771
	Context in Bare noun	44.14	9.03	4.89	<.001*	42.96	10.64	4.04	<.001*	3.53	13.25	0.27	.791
	Context in Colour-modified noun	-17.22	10.02	-1.72	.100	-17.92	7.93	-2.26	.030*	0.37	11.74	0.03	.975

Post-critical region (*was soaking*). First, first-pass regressions were 5.9% more likely following bare nouns in the two-referent context than in the one-referent context, whereas the referential context did not influence first-pass regressions following colour-modified nouns. Second, regression-path times were 75ms shorter following bare nouns in the one-referent context than in the two-referent context. In contrast, they were marginally (26ms) longer following colour-modified nouns in the one-referent context. Third, total times following bare nouns were 88ms longer in the two-referent context than in the one-referent context, whereas total times following colour-modified nouns were 39ms shorter in the two-referent context, demonstrating a cross-over interaction. First-pass times following neither bare nouns nor colour-modified nouns were affected by the referential context.

Final region (*on the floor*). We found 10.5% more first-pass regressions following bare nouns in the two-referent context than in the one-referent context, whereas first-pass regressions following colour-modified nouns were unaffected by the referential context. Also, regression-path times following bare nouns were 191ms longer in the two-referent context than in the one-referent context, whereas regression-path times following color-modified nouns did not reliably differ between the referential contexts.

Discussion.

As in Experiment 2, referential ambiguity delayed total times, not first-pass times, for the bare noun region. After reading past ambiguous bare nouns, readers were more likely to regress and spend more time interpreting earlier parts of the sentences, as shown by first-pass regressions and regression-path times in the post-critical and final regions. Hence, as in Experiment 2, referential ambiguity had an effect, albeit in later eye movement measures. Yet unlike in Experiment 2, where redundancy of size-modifiers immediately affected first-pass times for the modified noun region (as well as reading measures for subsequent regions), redundancy of color modifiers in the current

experiment only affected total times for the post-critical region; readers spent more time in reading the post-critical region when a colour-modifier occurred in a one-referent context and hence it was redundant. The absence of a first-pass effect on color-modified referring expressions provides support for the meaning-based early redundancy hypothesis, which predicts that the early redundancy effect is not due to the violation of Gricean maxims per se; it occurs only if readers fail to access a contrast set implied by the meaning of the modifier. Because the meaning of color does not require a contrast set, redundancy of a color modifier does not immediately slow down reading.

Experiment 4

The fourth and final experiment examined the meaning-based early redundancy effect more directly, by comparing the effect of redundancy on color and size modifiers in a single experiment. As shown in (7), the one-referent and two-referent context sentences were followed by either a color adjective (7a & 7b) or a size adjective (7c & 7d). If the early redundancy effect is driven by Gricean expectations about optimal informativeness, then readers should immediately adopt a contrastive interpretation for any adjective, regardless of their meaning. Hence, redundancy of both color and size adjectives should quickly disrupt comprehension. That is, following both size and color modifiers, first-pass times for the modified noun region should be longer in the one-referent context as compared to the two-referent context. Alternatively, the meaning-based early redundancy hypothesis predicts that the early effect of redundancy is driven by the meaning of the modifier. Hence, whereas first-pass times for size-modified nouns should be longer in the one-referent context than in the two-referent context, first-pass times for the color-modified nouns should be unaffected by the context. Redundancy of colour adjectives may affect comprehension later, such that the effect will emerge in regression measures for the following words, as was the case with the effect of ambiguity for bare nouns, since both effects may be driven by inference following the violation of Gricean maxims.

(7a) There was one towel on the floor. The white towel was soaking on the floor.

(7b) There were two towels on the floor. The white towel was soaking on the floor.

(7c) There was one towel on the floor. The small towel was soaking on the floor.

(7d) There were two towels on the floor. The small towel was soaking on the floor.

Method

Participants. Forty new participants (28 females) were recruited from the University of Strathclyde student community as before. The data from three participants were replaced with new participants due to too many track losses, an experimenter error, or a low accuracy rate in answering the comprehension questions (over 50% errors).

Materials. The context sentence set the referential context as in Experiments 2 and 3. Unlike in the previous experiments, the target sentence always started with a modified noun. In the color adjective conditions (7a, 7b), the noun was modified by a color adjective, whereas in the size adjective condition (7c, 7d), the noun was modified by a size adjective. The color and size adjectives were matched on length within each item. In total, there were 48 experimental items.

Design. There were four conditions in total: *adjective* (color vs. size) × *referential context* (one-referent vs. two-referent). The conditions were distributed across 4 lists as before. Ten participants were randomly assigned to each list.

Procedure. This was the same as in Experiments 1 and 2.

Results

The eye movement data were prepared as before and 5.0% of trials were excluded due to track losses, blinks and other errors in the critical region (4.2% in the color modifier/one-referent condition; 4.1% in the color modifier/two-referent condition, 6.6% in the size modifier/one-referent condition and 5.1% in the size modifier/two-referent condition). Participants answered the comprehension questions accurately (99%). As before, we analysed the reading time measures using

mixed effect models, whereby referential context (one referent vs. two referents), adjective (color adjective vs. size adjective) and the interaction between the two were included as fixed effects. Our primary concern was whether referential context differentially influenced the reading of color and size modifiers. Hence, we focused our analyses on the presence or absence of the context × adjective interaction and the effect of context in each modifier condition. Table 7 reports the means of the reading time measures for the target sentence, and Table 8 presents the results from each region.

Table 7. Means of eye movement measures by region in Experiment 4

		Region		
Expression	Context	Critical	Post-critical	Final
First-pass time (ms)				
Color-modified noun	One referent	424 (8)	333 (8)	402 (12)
Color-modified noun	Two referents	437 (9)	323 (7)	391 (10)
Size-modified noun	One referent	434 (9)	314 (7)	377 (10)
Size-modified noun	Two referents	411 (8)	317 (7)	408 (11)
First-pass regressions (%)				
Color-modified noun	One referent	2.4 (0.7)	14.2 (1.7)	73.4 (2.1)
Color-modified noun	Two referents	2.0 (0.6)	9.1 (1.4)	73.9 (2.1)
Size-modified noun	One referent	1.3 (0.5)	13.5 (1.6)	71.2 (2.2)
Size-modified noun	Two referents	3.1 (0.8)	10.0 (1.4)	71.0 (2.2)
Regression-path time (ms)				
Color-modified noun	One referent	440 (9)	392 (11)	966 (32)
Color-modified noun	Two referents	449 (10)	362 (9)	852 (25)
Size-modified noun	One referent	438 (9)	382 (11)	952 (32)
Size-modified noun	Two referents	430 (9)	366 (11)	932 (30)
Total time (ms)				
Color-modified noun	One referent	580 (14)	433 (10)	476 (14)
Color-modified noun	Two referents	559 (14)	425 (10)	451 (12)
Size-modified noun	One referent	600 (15)	447 (11)	469 (13)
Size-modified noun	Two referents	554 (13)	432 (11)	494 (14)

Note: standard errors in brackets.

Table 8. Analyses on eye movement measures by region in Experiment 4

		Critical region				*Post-critical region*				*Final region*			
		β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
First-pass time	Adjective	-4.27	4.31	-0.99	.329	-6.22	3.14	-1.98	.048*	-1.50	5.58	-0.27	.790
	Context	-2.47	3.52	-0.70	.491	-1.35	3.15	-0.43	.667	5.34	5.39	0.99	.327
	Adjective x Context	-7.94	3.45	-2.30	.027*	3.96	3.20	1.24	.222	10.28	4.65	2.21	.034*
	Context in Color modifiers	5.32	4.89	1.09	.284	-5.07	4.66	-1.09	.283	-5.54	6.89	-0.80	.431
	Context in Size modifiers	-10.93	5.07	-2.16	.043*	2.39	4.46	0.54	.593	15.82	6.50	2.43	.015*
		β	*SE*	*z*	*p*	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*
First-pass regressions	Adjective	-0.05	0.25	-0.19	.851	0.00	0.10	0.01	.994	-0.09	0.06	-1.35	.176
	Context	0.15	0.25	0.60	.547	-0.27	0.10	-2.65	.008*	0.02	0.08	0.23	.818
	Adjective x Context	0.25	0.26	0.98	.329	0.06	0.09	0.67	.504	-0.01	0.06	-0.22	.829
	Context in Color modifiers	-0.10	0.31	-0.33	.745	-0.30	0.15	-2.03	.043*	0.03	0.09	0.34	.732
	Context in Size modifiers	0.37	0.39	0.95	.341	-0.21	0.12	-1.69	.091	0.00	0.09	-0.04	.966
		β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Regression-path time	Adjective	-4.87	5.27	-0.92	.362	-0.78	5.02	-0.16	.879	18.83	12.64	1.49	.145
	Context	0.72	3.79	0.19	.851	-11.06	5.80	-1.91	.064	-26.83	14.10	-1.90	.070
	Adjective x Context	-2.36	3.71	-0.64	.523	4.27	5.53	0.77	.444	19.40	12.52	1.55	.121
	Context in Color modifiers	3.04	5.20	0.58	.559	-15.35	7.18	-2.14	.042*	-49.23	21.68	-2.27	.031*
	Context in Size modifiers	-2.12	5.63	-0.38	.708	-7.57	7.45	-1.02	.315	-6.00	18.12	-0.33	.741
		β	*SE*	*t*	*p*	β	*SE*	*t*	*p*	β	*SE*	*t*	*p*
Total time	Adjective	2.84	6.06	0.47	.641	5.69	5.20	1.10	.279	9.65	7.09	1.36	.184
	Context	-17.65	6.78	-2.60	.015*	-4.51	5.26	-0.86	.397	1.94	5.92	0.33	.745
	Adjective x Context	-5.48	5.92	-0.93	.355	-1.67	5.36	-0.31	.758	10.99	5.49	2.00	.046*
	Context in Color modifiers	-12.98	10.42	-1.25	.222	-3.15	6.95	-0.45	.654	-9.85	7.91	-1.25	.219
	Context in Size modifiers	-23.57	8.46	-2.79	.005*	-6.65	7.07	-0.94	.347	13.50	7.86	1.72	.086

Critical region (the white towel/the small towel). First-pass times showed a significant adjective × context interaction: The referential context did not significantly influence the first-pass times for color-modified nouns, whereas size-modified nouns were read 23ms more slowly in the one-referent than in the two-referent condition. In total times, the reading times for size-modified descriptions were 46ms longer in the one-referent than in the two-referent context. In contrast, no reliable context effect was found with color-modified descriptions, though there was no significant context × adjective interaction in total times. No other context effects were found in this region. Across all measures, there were no significant main effects of adjective.

Post-critical region (was soaking). In both adjective conditions, there were more regressions in the one-referent than the two-referent context; the effect was fully significant following color modifiers but marginal following size modifiers. Second, following color adjectives, regression-path times were 30ms longer in the one-referent condition than the two-referent condition, whilst no significant difference was found following size adjectives, though the adjective × context was not significant.

Final region (on the floor). First, there was a significant adjective × context interaction in the first-pass times: Whereas first-pass times were unaffected by the referential context following color-adjectives, they were 31ms longer in the one-referent than in the two-referent context following size adjectives. Second, following color adjectives, regression-path times were 114ms longer in the one-referent than the two-referent condition, whereas no reliable context effect was found following size adjectives, though the adjective × context interaction was not significant. Finally, total times showed a significant adjective × context interaction, though the effect of referential context was not fully significant following either color or size adjectives.

Discussion

As in Experiments 1 and 2, first-pass times for the noun phrase region showed that size-modified noun phrases took longer to read in the one-referent context than in the two-referent context. Total times in the same region also showed longer reading times for size-modified nouns when they were redundant in the context. Crucially, as in Experiment 3, the referential context did not significantly affect reading times for color-modified nouns in the noun phrase region; if anything, first-pass times for the color-modified nouns were numerically longer in the two-referent context. The redundancy of the color adjectives only affected regression measures in the post-critical and final regions, with more regressions in the post-critical region and longer regression-path times in the post-critical and the final regions following a one-referent than a two-referent context. Hence, the effect of redundant color adjectives was delayed relative to that of redundant size adjectives, providing direct support for the meaning-based hypothesis that an early redundancy effect only occurs when the lexical semantics of the modifier sets up a referential contrast that is absent in the referential context.

General Discussion

We began this article by presenting three hypotheses concerning how referential ambiguity and referential redundancy might disrupt online reading. Some research on parsing (e.g., Altmann & Steedman, 1988; Crain & Steedman, 1985; Tanenhaus et al., 1995; Van Berkum et al., 1999) has shown that referential context immediately affects parsing, as comprehenders adopt analyses that avoid ambiguity as well as redundancy. The first hypothesis, which we termed the *strong version* of the Gricean hypothesis, therefore predicted that any violation of the Gricean maxims rapidly disrupts comprehension, both when readers encounter referential ambiguity and referential redundancy. In contrast, off-line rating studies (Davies & Katsos, 2013; Engelhardt et al., 2006) indicated that language users tend to consider ambiguous descriptions more problematic than redundant descriptions. Hence, the *ambiguity-first* hypothesis proposed that ambiguity disrupts comprehension

more severely or perhaps more quickly than redundancy. Finally, based on research that showed that comprehenders are less sensitive to referential ambiguity, the *redundancy-first* hypothesis predicted that redundancy affects comprehenders more quickly than ambiguity.

Experiments 1-2 showed that readers immediately slowed down when encountering redundant size-modifiers, as indicated by longer first-pass times for the modified noun region in one- than two-referent contexts. In contrast, ambiguous bare nouns did not immediately influence reading behaviour; in Experiments 2 and 3, ambiguity affected total times for the bare noun region and regression-based measures in the regions following the referring expression, while there was no clear ambiguity effect in Experiment 1. These findings provide evidence against the ambiguity-first hypothesis as well as the strong version of the Gricean view, which assumes that both ambiguity and redundancy similarly disrupt early comprehension processes. Crucially, Experiments 3 and 4 showed that the early redundancy effect was dependent on the meaning of the modifier: Unlike size modifiers, the redundancy of color modifiers did not immediately slow down reading. This indicates that initial referential processing is led by the lexico-semantic representation of the referring expression, rather than Gricean expectations about optimal informativeness: When size-modified definite noun phrases follow a context with only one referent compatible with the head noun, the processor immediately slows down, because the semantic processing of the modifier requires a contrast set. In contrast, the semantic processing of color modifiers does not involve comparison with other discourse entities, so their redundancy did not result in immediate difficulty.

Our results appear to be in contrast with studies that showed that redundancy facilitates referent identification. Arts et al. (2011) reported faster referent identification times after redundant reference for some specific combinations of modifiers. In their study, participants read a written description presented on a computer screen, before being shown a visual display of different objects. The identification times included the time taken for participants to identify the referent and memorize the number given below each object, but not the time for processing the referring

expression itself, which was presented separately from the visual display. Also, in their study, overspecified descriptions may have been highly informative for comprehenders, as they contained perceptually more salient properties (e.g., *the large round white button*) than more concise descriptions (e.g., *the round button*). When more concise descriptions contained highly effective perceptual cues such as location (e.g., *the button at the top left*), then inclusion of size (*the large button at the top*) or size and color (*the large white button at the top*) resulted in *longer* reaction times.

Our findings may also seem inconsistent with Engelhardt et al. (2011). In their study, participants took longer to identify the target object following redundancy of both size and color adjectives, though their behavioural data showed a larger redundancy effect with size adjectives, and there was no significant adjective × context interaction in an N400-like effect that occurred at around 450-570ms following redundant adjective onset. Other research also suggests that comprehenders immediately interpret modifiers contrastively even if the modifiers are not inherently contrastive. In Hanna et al. (2003), participants listened to and carried out instructions such as *Pick up the empty …* when the visual array contained two martini glasses, one empty and one full of olives, and an empty long-drink glass. They found that on hearing *empty*, participants were more likely to look at the empty martini glass than at the empty normal glass, indicating that listeners contrastively interpreted the adjective *empty*, even though it is not an inherently comparative modifier; whether the object is empty or not is independent of the state of another object. Moreover, Sedivy et al. (1999) showed, in another visual world experiment, that participants rapidly interpret non-comparative adjectives with denotational properties such as colour, material, or shape of an object contrastively. Following *Touch the pink comb*, participants heard *Now touch the yellow comb/bowl*. Participants were quicker at fixating the target object in the second instruction when the object was a comb rather than a bowl, that is, when the adjective contrasted objects of the same type (two different combs). However, Ferreira, Foucart and Engelhardt (2013) found that when the visual display contained contrasting

objects, listeners tended to expect that speakers would refer to one of the contrasting objects *even before they heard the modifier*. Hence, the immediate referential contrast effects observed in Sedivy (1999) and Hanna et al. (2003) might reflect comprehenders' non-linguistic expectations about the likely referent rather than their interpretation of the modifier itself. Another possibility is that the goal of the referential communication task in visual world studies favoured a contrastive interpretation, as it may have enhanced listeners' Gricean expectation that cooperative speakers should use modifiers to help them discriminate contrasting referents, rather than attributing redundant information for the task at hand. Either way, we should note that in the current study, redundancy of colour modifiers did have an effect in later regions, indicating that comprehenders do adopt a contrastive interpretation for these modifiers, but not as quickly as for size modifiers.

Some researchers have argued that language comprehenders do not always build fully specified representations, unless the task requires the processor to do so (Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Ferreira, 2003; Swets, Desmet, Clifton, & Ferreira, 2008). Moreover, Piantadosi, Tily, and Gibson (2012) argued that ambiguity arises for the sake of communicative efficiency; ambiguous words are usually short and frequent and hence easy to produce or comprehend. Often the context provides sufficient cues about the intended meaning, and crucially, comprehenders "actively use context" to resolve ambiguity, as shown by studies on anaphor resolution (e.g., *Sally frightened Mary because she is very timid*, Garvey & Caramazza, 1974; Stevenson, Crawley, & Kleinman, 1994). Moreover, research has shown that comprehenders do not always resolve ambiguous pronouns (e.g., Greene et al., 1992; Levine, Guzman, & Klin, 2000) or they "underspecify" ambiguous pronouns until further disambiguating information is encountered (Stewart et al., 2007; cf. Swets et al., 2008). This may be the reason why ambiguity did not slow down reading immediately in our study, because readers initially accommodated ambiguity, anticipating disambiguating information later in the sentence. They then made regressive eye movements from subsequent regions when they realized that there was no disambiguating

information. Although these accounts are consistent with our data, our account goes beyond previous accounts: Assuming that referential difficulty arises immediately when comprehenders fail to identify at least one referent relevant for the meaning of the referring expression, our account explains *why* redundant modifiers impair comprehension immediately when the referential context lacks the referential contrast implied by the modifier and *why* ambiguous bare nouns do not disrupt comprehension initially.

Finally, some research suggests that redundancy has its own communicative function. Vonk, Hustinx, and Simons (1992), for instance, argued that the use of overly specific referring expressions at episodic boundaries signals a topic shift. In their experiment, participants were more likely to use names or definite descriptions (relative to pronouns) to refer to characters after a shift in time and place (see also Anderson, Garrod, & Sanford, 1983, for a similar finding). However, these studies did not investigate whether overly specific referring expressions nevertheless slowed down comprehension. The current study showed that in relatively confined neutral contexts, overly explicit expressions do pose difficulty when the redundant adjective implies a contrast set. Future research may wish to examine whether the cost of referential redundancy is modulated by the discourse functions of the referring expressions and whether such modulation takes place immediately.

In conclusion, our findings speak against the view that Gricean pragmatic expectations about optimal informativeness have an immediate impact on reading, commonly held in research on parsing. Referential processing is highly incremental, in that as soon as comprehenders encounter the referring expression, they launch a search for the relevant discourse referents. Crucially, this initial referential processing is primarily led by the lexico-semantic representation of the referring expression; as a result, ambiguous descriptions that allow the identification of at least one semantically compatible referent do not immediately slow down reading, whilst modifiers that semantically imply a contrast impair comprehension immediately when they occur in a context with no referential alternative. Violations of the Gricean maxims have an effect only after referential

processing has identified relevant discourse referents on the basis of the meaning of the referring expression.

Acknowledgements

We thank Gemma Roddie and Rachel Dickie for their help with material construction, and Jaime Armit, Eve Norris, Elle Gemmell, Julie Fuglsang, Vladislava Segen, Lindsay Holton, Scott Orr and Maria Nella Carminati for their help with data collection. Kumiko Fukumura acknowledges support from Experimental Psychology Society Grant 2015 and Leverhulme Trust Research Project Grant RPG-2016-253.

References

Altmann, G. T. M., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30*, 191–238.

Altmann, G. T. M., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language, 31*, 685–712.

Altmann, G. T. M., Garnham, A., & Henstra, J.-A. (1994). Effects of syntax in human sentence parsing: Evidence against a structure-based proposal mechanism. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 209–216.

Anderson, A., Garrod, S. C., & Sanford, A. J. (1983). The accessibility of pronominal antecedents as a function of episode shifts in narrative text. *Quarterly Journal of Experimental Psychology Section a-Human Experimental Psychology, 35*, 427-440.

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

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390-412.

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.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. arXiv preprint arXiv:1506.04967.

Bates, B., Mächler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1-48.

Belke, E., & Meyer, A. S. (2002). Tracking the time course of multidimensional stimulus discrimination: Analyses of viewing patterns and processing times in ‘‘same’’-‘‘different’’ decisions. *European Journal of Cognitive Psychology*, *14*, 237-266.

Bierwisch, M. (1987). The semantics of gradation. In M. Bierwisch & E. Lang (Eds.), *Dimensional Adjectives* (pp.471-514). Berlin: Springer-Verlag.

Britt, M.A. (1994). The interaction of referential ambiguity and argument structure in the parsing of prepositional phrases. *Journal of Memory and Language, 33,* 251-283.

Chambers, C.G., Tanenhaus, M.K., & Magnuson, J.S. (2004). Action-based affordances and syntactic ambiguity resolution. *Journal of Experimental Psychology: Learning, Memory & Cognition, 30,* 687-696.

Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42,* 368–407.

Clifton, C. & Ferreira, F. (1989). Ambiguity in context. *Language and Cognitive Processes, 4,* 77-104.

Cook, A. E., & O'Brien, E. J. (2014). Knowledge activation, integration, and validation during narrative text comprehension. *Discourse Processes, 51,* 26-49.

Crain, S., & Steedman, M. (1985). On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Karttunnen, & A. Zwicky (Eds.), *Natural language parsing* (pp. 320–358). Cambridge, UK: Cambridge University Press.

Danks, J. H., and Glucksberg, S. 1971. Psychological scaling of adjective orders. *Journal of Verbal Learning and Verbal Behavior 10,* 63-67

Davies, C. & Katsos, N. (2013). Are speakers and listeners 'only moderately Gricean'? A response to Engelhardt, Bailey and Ferreira (2006).*Journal of Pragmatics, 49,* 78-106.

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

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

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47,* 164–

203.

Ferreira, F., & Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348–368.

Ferreira, F., Foucart, A., & Engelhardt, P. E. (2013). Language processing in the visual world: Effects of preview, visual complexity, and prediction, *Journal of Memory and Language*, *69,* 165-182.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 559-586). Hillsdale, NJ: Lawrence Erlbaum Associates.

Garnham, A., Oakhill, J., & Cruttenden, H. (1992).The role of implicit causality and gender cue on the interpretation of pronouns. *Language and Cognitive Processes*, *7,* 231–255.

Garrod, S., O'Brien, E.J., Morris, R.K., & Rayner, K. (1990). Elaborative inferencing as an active or passive process. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 250-257.

Garrod, S., & Terras, M. (2000). The contribution of lexical and situational knowledge to resolving discourse roles: bonding and resolution. *Journal of Memory and Language*, *42*, 526–544.

Garvey, C., & Caramazza, A. (1974). Implicit causality in verbs. *Linguistic Inquiry, 5*, 459-464.

Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition, 32*, 99–156.

Greene, S., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 266–283.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics: Vol. III. Speech acts* (pp. 41–58). New York: Academic Press.

Hanna, J.E., Tanenhaus, M.K., & Trueswell, J.C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, *49,* 43-61.

Hintzman, D. L. (1986). "Schema abstraction"' in a multiple-trace memory model. *Psychological Review, 93,* 411-428.

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.

Levine, W. H., Guzman, A. E., & Klin, C. M. (2000). When anaphor resolution fails. *Journal of Memory and Language, 43,* 594–617.

Logačev, P., & Vasishth, S. (2016). A multiple-channel model of task-dependent ambiguity resolution in sentence comprehension. *Cognitive Science, 40,* 266-298.

Martin, J. E. (1969a). Semantic determinants of preferred adjective order. *Journal of Verbal Learning and Verbal Behavior, 8,* 697-704.

Martin, J. E. (1969b). Some competence-process relationships in noun phrases with prenominal and postnominal adjectives. *Journal of Verbal Learning and Verbal Behavior, 8,* 471-480.

Murray, W. S., & Liversedge, S. P. (1994). Referential context effects on syntactic processing. In C. Clifton, Jr., L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp.359-388). Hillsdale, NJ: Lawrence ErlbaumAssociates, Inc.

Neale, S. (1990). *Descriptions*. Cambridge, MA: The MIT Press.

Nieuwland, M.S., Otten, M. & Van Berkum, J.J.A. (2007). Who are you talking about? Tracking discourse-level referential processes with ERPs. *Journal of Cognitive Neuroscience, 19,* 1-9.

Olson, D. R. (1970). Language and thought - aspects of a cognitive theory of semantics. *Psychological Review, 77,* 257-273.

Osgood, C. E. (1971). Where do sentences come from? In D. D. Steinberg & L. A. Jakobovits (Eds.), *Semantics: An interdisciplinary reader in philosophy, linguistics and psychology* (pp. 497–529). London: Cambridge University Press.

Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics, 27,* 89-110.

Piantadosi, S.T., Tily, H. & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition, 122*, 280-291.

Piantadosi, S.T., Tily, H. & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition, 122*, 280-291.

Poesio, M. (2003). Associative descriptions and salience: a preliminary investigation. In *Proceedings of the ACL Workshop on Anaphora* (pp. 31-38), Budapest.

Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical Pragmatics* (pp. 223-255). New York: Academic Press.

R Core Team (2016). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59-108.

Rigalleau, F., & Caplan, D. (2000). Effects of gender marking in pronominal coindexation. *Quarterly Journal of Experimental Psychology, 53A*, 23–52.

Rigalleau, F., Caplan, D., & Baudiffier, V. (2004). New arguments in favour of an automatic gender pronominal process. *Quarterly Journal of Experimental Psychology*, *57A,* 893–933.

Russell, B. (1905). On denoting. *Mind, 14*, 479-493.

Schriefers, H., & Pechmann, T. (1988). Incremental production of noun-phrases by human speakers. In M. Zock & G. Sabah (Eds.), Advances in natural language generation (pp. 172-179). London: Pinter.

Sedivy, J. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of Psycholinguistic Research, 32,* 3-23.

Sedivy, J., Tanenhaus, M., Chambers, C., & Carlson, G. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition, 71*, 109–147.

Spivey, M. J., & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modelling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1521–1543.

Spivey, M., Tanenhaus, M., Eberhard, K., & Sedivy, J. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, *45*, 447-481.

Steedman, M., & Altmann, G. (1989). Ambiguity in context: A reply. *Language and Cognitive Processes, 4*, 105–122.

Stevenson, R. J., Crawley, R. A., & Kleinman, D. (1994). Thematic roles, focus and the representation of events. *Language and Cognitive Processes, 9*, 519-548.

Stewart, A. J., Holler, J., & Kidd, E. (2007). Shallow processing of ambiguous pronouns: Evidence for delay. *Quarterly Journal of Experimental Psychology, 60*, 1680-1696.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition*, *36*, 201–216.

Tanenhaus, M., Spivey-Knowlton, M., Eberhard, K., & Sedivy, J. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632.

Traxler, M. J., Pickering, M. J., & Clifton, C., Jr. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, *39*, 558–592.

Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, *41*, 147–182.

Van Deemter, K. (2002). Generating referring expressions: Boolean extensions of the incremental algorithm. *Computational Linguistics, 28*, 37-52.

Van Gompel, R. P. G., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language, 45,* 225–258.

Vonk, W., Hustinx, L., & Simons, W. H. G. (1992). The use of referential expressions in structuring discourse. *Language and Cognitive Processes, 7,* 301-333.

Zagar, D., Pynte, J., & Rativeau, S. (1997). Evidence for early closure attachment on first pass reading times in French. *Quaterly Journal of Experimental Psychology, 50A,* 421-438.

Appendix: Materials

Experiments 1 and 2

In Experiment 1, the noun before the slash was used in the one-referent context conditions and the noun after the slash in the two-referent conditions. The adjective in brackets was used in the modified noun conditions, whereas the bare noun conditions did not have this adjective. The order of the two noun phrases in the context sentence was reversed in additional counterbalancing conditions. In Experiment 2, the context sentence contained the numeral *one* (one-referent condition) or *two* (two-referent condition), e.g., There was one button/two buttons on the machine. The target sentence was the same as in Experiment 1.

1. There was a big button and a little lever/button on the machine. The (big) button was to reset the machine.
2. There was a little shell and a big pebble/shell in the sand. The (little) shell was taken for the shop.
3. There was a small cup and a big bowl/cup on the table. The (small) cup was knocked off the side.
4. There was a large aeroplane and a small helicopter/aeroplane above the city. The (large) aeroplane was flying for many hours.
5. There was a small ring and a large necklace/ring in the jewellery box. The (small) ring was stolen by the thief.
6. There was a large mouse and a small squirrel/mouse in the house. The (large) mouse was caught in the trap.
7. There was a big slide and a small swing/slide in the park. The (big) slide was always most popular.
8. There was a small towel and a large robe/towel in the bathroom. The (small) towel was soaking on the floor.

9. There was a small diagram and a big graph/diagram in the report. The (small) diagram only confused people more.
10. There was a large dog and a small cat/dog in the kitchen. The (large) dog bolted out the door.
11. There was a little ladybird and a big spider/ladybird on the plant pot. The (little) ladybird fell off the plant pot.
12. There was a big tattoo and a little piercing/tattoo on the woman's neck. The (big) tattoo attracted attention.
13. There was a small scarf and a big jacket/scarf on the back of the chair. The (small) scarf looked quite old and worn.
14. There was a large wasp and a small butterfly/wasp in the garden. The (large) wasp flew into the house.
15. There was a little candle and a large oil-burner/candle on the dressing table. The (little) candle had never been lit before.
16. There was a large house and a small shed/house in the street. The (large) house had remained abandoned.
17. There was a small sofa and a large chair/sofa in the waiting room. The (small) sofa was filled with cushions.
18. There was a little box and a big bag/box at the party. The (little) box contained fancy gifts.
19. There was a small monkey and a large giraffe/monkey in the zoo. The (small) monkey was the main attraction.
20. There was a small leopard and a big lion/leopard in the zoo. The (small) leopard was clearly very hungry.
21. There was a large screwdriver and a small hammer/screwdriver in the box. The (large) screwdriver was essential for every job.

22. There was a small ball and a big hula-hoop/ball in the toy store. The (small) ball was covered in red dots.
23. There was a big bin and a small mailbox/bin outside the house. The (big) bin was overflowing after a holiday.
24. There was a little painting and a big mirror/painting on the wall. The (little) painting was taking pride of place.
25. There was a big mall and a small boutique/mall on the street. The (big) mall was closed down recently.
26. There was a little balloon and a big cake/balloon at the party. The (little) balloon was given by a friend.
27. There was a large cow and a small pig/cow on the farm. The (large) cow was lying in the field.
28. There was a small guitar and a big keyboard/guitar on the stage. The (small) guitar was played by the singer.
29. There was a little ladder and a big plank/ladder against the wall. The (little) ladder was left by the builders.
30. There was a big banana and a small grapefruit/banana on the kitchen counter. The (big) banana was put into the blender.
31. There was a little tower and a big church/tower in the village. The (little) tower was popular with tourists.
32. There was a large medal and a small trophy/medal in the cabinet. The (large) medal was won in a local rally.
33. There was a little cut and a big bruise/cut on the man's arm. The (little) cut was caused by falling.
34. There was a big hotdog and a small burger/hotdog on the plate. The (big) hotdog was covered in mustard.

35. There was a large key and a small padlock/key on the table. The (large) key was for the school locker.
36. There was a big pizza and a small lasagne/pizza in the buffet. The (big) pizza was divided into six sections.
37. There was a small notepad and a big jotter/notepad on the desk. The (small) notepad contained the schedule.
38. There was a big iguana and a small hamster/iguana in the pet shop. The (big) iguana ran around in circles.
39. There was a big pie and a small biscuit/pie on the shelf. The (big) pie was covered with cream.
40. There was a little bird and a big hedgehog/bird under the tree. The (little) bird was eating from the feeder.
41. There was a large alligator and a small hippo/alligator in the swamp. The (large) alligator was hiding under a log.
42. There was a big camera and a small phone/camera in the drawer. The (big) camera hadn't been used in years.
43. There was a little bowtie and a big hat/bowtie on the bed. The (little) bowtie wasn't worn every day.
44. There was a big pen and a small eraser/pen in the pencil case. The (big) pen still looked brand new.
45. There was a little flask and a big jug/flask in the basket. The (little) flask contained tea for later.
46. There was a large puddle and a small cone/puddle on the road. The (large) puddle caused cars to swerve.
47. There was a large star and a small circle/star on the birthday card. The (large) star was covered in glitter.

48. There was a small hairdryer and a large straightener/hairdryer on the desk. The (small) hairdryer was making a loud noise.

Experiments 3 and 4

The one-referent context conditions contained the numeral *one*, whereas the two-referent conditions contained *two*. In Experiment 4, the adjective before the slash occurred in the color-modifier conditions and the adjective after it occurred in the size-modifier conditions. Experiment 3 compared the color-modifier conditions with bare noun conditions (conditions without adjective).

1. There was one button/There were two buttons on the machine. The red/big button was to reset the machine.
2. There was one shell/There were two shells in the sand. The yellow/little shell was taken for the shop.
3. There was one cup/There were two cups on the table. The green/small cup was knocked off the side.
4. There was one helicopter/There were two helicopters above the city. The white/large helicopter was flying for many hours.
5. There was one ring/There were two rings in the jewellery box. The purple/little ring was stolen by the thief.
6. There was one mouse/There were two mice in the house. The white/large mouse was caught in the trap.
7. There was one slide/There were two slides in the park. The red/big slide was always very popular.
8. There was one towel/There were two towels in the bathroom. The white/small towel was soaking on the floor.

9. There was one clock/There were two clocks in the cafe. The green/small clock was five minutes late.
10. There was one dog/There were two dogs in the kitchen. The brown/large dog bolted out the door.
11. There was one beetle/There were two beetles on the plant pot. The yellow/little beetle fell off the plant pot.
12. There was one tattoo/There were two tattoos on the woman's neck. The green/large tattoo attracted attention.
13. There was one scarf/There were two scarves on the back of the chair. The black/small scarf looked quite old and worn.
14. There was one butterfly/There were two butterflies in the garden. The white/large butterfly flew into the house.
15. There was one candle/There were two candles on the dressing table. The yellow/little candle had never been lit before.
16. There was one house/There were two houses in the street. The white/large house had remained abandoned.
17. There was one sofa/There were two sofas in the waiting room. The black/small sofa was filled with cushions.
18. There was one box/There were two boxes at the party. The yellow/little box contained fancy gifts.
19. There was one monkey/There were two monkeys in the zoo. The black/small monkey was the main attraction.
20. There was one turtle/There were two turtles in the pond. The green/large turtle was over 90 years old.

21. There was one screwdriver/There were two screwdrivers in the toolbox. The black/large screwdriver was essential for every job.
22. There was one ball/There were two balls in the toy store. The white/small ball was quite expensive.
23. There was one bin/There were two bins outside the house. The green/large bin was full after the holiday.
24. There was one vase/There were two vases on the windowsill. The orange/little vase was full of flowers.
25. There was one lamp/There were two lamps in the bedroom. The green/small lamp gave off a bright light.
26. There was one balloon/There were two balloons at the party. The green/little balloon was given by a friend.
27. There was one cow/ There were two cows on the farm. The brown/large cow was lying in the field.
28. There was one guitar/There were two guitars on the stage. The white/small guitar was played by the singer.
29. There was one ladder/There were two ladders against the wall. The black/large ladder was left by the builders.
30. There was one apple/There were two apples on the kitchen counter. The green/small apple was put into the blender.
31. There was one tower/There were two towers in the village. The white/small tower was popular with tourists.
32. There was one plate/There were two plates in the cabinet. The green/large plate was apparently very precious.

33. There was one cat/There were two cats sitting in the sun. The black/large cat seemed to be asleep.
34. There was one chocolate/There were two chocolates in the box. The white/large chocolate had melted completely.
35. There was one key/There were two keys on the table. The black/large key was for the school locker.
36. There was one tub/There were two tubs in the lunchbox. The white/small tub contained fruit salad.
37. There was one notepad/There were two notepads on the desk. The green/small notepad contained the schedule.
38. There was one iguana/There were two iguanas in the pet shop. The yellow/little iguana ran around in circles.
39. There was one book/There were two books on the shelf. The brown/large book was left by the boy.
40. There was one bird/There were two birds under the tree. The yellow/little bird was eating from the feeder.
41. There was one alligator/There were two alligators in the swamp. The brown/large alligator was hiding under a log.
42. There was one camera/There were two cameras in the drawer. The red/big camera hadn't been used in years.
43. There was one bowtie/There were two bowties on the bed. The yellow/little bowtie wasn't worn very often.
44. There was one pen/There were two pens in the pencil case. The red/big pen still looked brand new.

45. There was one flask/There were two flasks in the basket. The orange/little flask contained tea for later.
46. There was one car/There were two cars parked in the garage. The white/small car had a number of dents.
47. There was one star/There were two stars on the birthday card. The green/large star was covered in glitter.
48. There was one hairdryer/There were two hairdryers on the desk. The white/small hairdryer was making a loud noise.