2008

The Effect of Context on the Interpretation of Noun-Noun Combinations: Eye Movement and Behavioral Evidence

Tony Mccaffrey
University of Massachusetts Amherst, amccaffr@psych.umass.edu

THE EFFECT OF CONTEXT ON THE INTERPRETATION OF NOUN-NOUN COMBINATIONS: EYE MOVEMENT AND BEHAVIORAL EVIDENCE

A Master's Thesis

by

TONY MCCAFFREY

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE

May 2008

Psychology

THE EFFECT OF CONTEXT ON THE INTERPRETATION OF NOUN-NOUN COMBINATIONS: EYE MOVEMENT AND BEHAVIORAL EVIDENCE

A Master's Thesis

by

TONY MCCAFFREY

Approved as to style and content by:

_____
Andrew Cohen, Chair

_____
Keith Rayner, Member

_____
Lisa Sanders, Member

_____
Lisa Scott, Member

_____
Melinda A. Novak, Department Chair
Psychology Department

ABSTRACT

THE EFFECT OF CONTEXT ON THE INTERPRETATION

OF NOUN-NOUN COMBINATIONS:

EYE MOVEMENT AND BEHAVIORAL EVIDENCE

MAY 2008

TONY MCCAFFREY, B.A., KNOX COLLEGE, GALESBURG, IL

M.S., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Andrew Cohen

Two experiments examined processing of "near-nonsense" noun-noun combinations (e.g., *dictionary treatment, olive signals*). In the first experiment, readers' eye movements were monitored as they read sentences containing such combinations, or control sentences containing easy-to-interpret adjective-noun combinations. A preceding context sentence either did or did not support a specific interpretation of the critical noun-noun combination. The earliest measures of processing difficulty were not modulated by the context manipulation, but on later measures, the potentially helpful context did alleviate difficulty. In the second experiment, participants provided detailed interpretations of the critical combinations, with and without the potentially helpful context sentence; the results confirmed that the context sentences encouraged specific interpretations of these combinations. The results suggest that a noun-noun combination is initially interpreted without taking into account the immediate context, but that this context may ultimately play a critical role.

CONTENTS

Page

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Noun-noun combinations are a common way to provoke new meanings from existing concepts.[1] For example, the combination *doll smile* might suggest either a smile on a doll or a smile on a child receiving a doll as a gift (Gerrig & Bortfeld, 1999). The study of noun-noun combinations informs research on both concepts and language comprehension. While concept research predominantly focuses on concepts in isolation (Murphy, 2004), more complex concepts can be built from combinations of simpler ones. A two-word phrase, such as a noun-noun combination, is a minimal linguistic construct for combining two concepts. Thus, the study of noun-noun combinations provides an entry point into the comprehension of larger linguistic structures such as phrases and sentences.

Although noun-noun combinations are usually studied in isolation, they are often encountered in a context that helps determine their meaning (Gagné & Spalding, 2004). When asked to construct a meaning for *doll smile* in isolation, for example, participants favor the meaning "a smile on a doll" (Gerrig & Bortfeld, 1999). However, encountering the combination in context can shift the favored meaning. Consider the following example:

> Aunt Bev's gifts to her niece and nephew elicited smiles. Bev says, "I saw a doll smile that was very impressive" (Gerrig & Bortfeld, 1999).

The first sentence acts as a context in which to interpret *doll smile* and clearly suggests a different meaning, perhaps a child's smile after receiving a doll.

The central issue addressed in the current research is the relation between the processes involved when meaning is derived from a noun-noun combination in isolation, and those involved when a noun-noun combination appears in a potentially relevant context. Two contrasting views on this question can be identified in the existing literature. According to the *independence* view (Gerrig & Bortfeld, 1999), fundamentally distinct interpretive processes are involved when noun-noun combinations are encountered in and out of context. Gerrig and Bortfeld created pairs of noun-noun combinations (e.g., *doll smile* and *baseball smile*) that had four distinct features. First, the second noun (i.e., *smile*) was the same for each member of the pair. Second, in isolation, one was easier to interpret than the other (e.g., *doll smile* was easier than *baseball smile*). Third, in the proper context, the members of the pair yielded comparable meanings (e.g., "smile on a child receiving a doll/baseball"). Fourth, for the easier member of the pair (i.e., *doll smile*), the meaning preferred in isolation differed from the meaning in context (i.e., "smile on a doll" and "smile on a child receiving a doll," respectively). Gerrig and Bortfeld (1999) recorded the time it took participants to read and respond to short discourses containing the critical combinations, and found that the proper context eliminated the disparity between the paired combinations; although it is easier to create a meaning for *doll smile* in isolation, in context, *doll smile* and *baseball smile* were of comparable difficulty. Because there was no indication that the preferred meaning for *doll smile* in isolation (i.e., "smile on a doll") interfered with the construction of a different meaning in context (i.e., "smile on a child receiving a doll"), Gerrig and

Bortfeld (1999) argued that creating a meaning in isolation is a different process from creating one in context.

Gagné and Spalding (2004) proposed that, for both theoretical and empirical reasons, this conclusion of independence was unwarranted. The stimuli from Gagné and Spalding's key experiment (Experiment 2) were short discourses that included noun-noun combinations in a neutral or supportive context, such as:

> Bob went to the magazine store to do some shopping. Bob had a busy week ahead of him. *First, he needed to finish his taxes, and then he needed to plan his trip to the mountains.* He needed to find some up-to-date information to help with his tasks. When Bob got to the store, he realized that he only had enough money to buy one magazine. After much deliberation, Bob decided to get a tax/mountain magazine.

The final sentence included just one of the target combinations, e.g., either *tax magazine* or *mountain magazine*. The neutral context did not include the italicized sentence. Thus, while the first sentence in both the neutral and supportive contexts primes *magazine*, only the supportive context provides information about what kind of magazine Bob might buy.

The two combinations in the final sentence were selected to be of differing difficulty when encountered in isolation; for this example, participants found it much easier to assign meaning to *tax magazine* than to *mountain magazine*. Gagné and Spalding interpreted this difference in difficulty in terms of relation-linking theory. Relation-linking theory states that, given a noun-noun combination, meaning construction is triggered by assuming a particular type of relation between the two nouns (Downing, 1977; Shoben, 1991). Different relation-linking theories propose between ten (Levi,

1978) and 15 (Gagné & Shoben, 1997) potential relations. Using the relation *about*, for example, to combine *tax* and *mountain* with *magazine* results in the meanings "a magazine about taxes" and "a magazine about mountains," respectively. The critical difference is that the relation *about* is used much more frequently with *tax* than it is with *mountain*. Gagné and Shoben (1997) have shown that it is easier to interpret an isolated combination when the relation frequency is high, so *tax magazine* is easier to interpret in isolation than *mountain magazine*.

Gagné and Spalding (2004) found that even when the meanings of both combinations were primed by context, as in the supportive context condition, *tax magazine* is still easier to interpret than *mountain magazine*. Gagné and Spalding concluded that relation frequency exerts an influence even when a combination is presented in a supportive context. These results led Gagné and Spalding to favor the *interdependence* view: the same processes are used to interpret both isolated and in-context combinations. In this view, context provides additional information that can be used during the creation of meaning, but does not fundamentally alter the process of interpreting a noun-noun combination.

The distinction between the independence and interdependence views is not as straightforward as it may appear. Gerrig and Bortfeld's results do not necessarily imply the existence of two independent processes. As Gagné and Spalding point out, a lack of direct interference between two possible meanings of *doll smile* does not necessarily mean that two independent processes are in play. For example, there could be a parallel process with two meanings racing for completion. In the supportive context, the less accessible meaning (i.e., "smile on a child's face after receiving a doll") may finish first

over the more accessible meaning (i.e., "smile on a doll"). On the other hand, Gagné and Spalding's results do not necessarily imply that a single process is active in both cases. The process of finding a reference for *mountain magazine* in a preceding discourse could be a different process than creating a meaning in isolation for this combination, even if the time-courses are comparable; more difficult combinations in isolation could also be more difficult to find a reference for, although for possibly different reasons.

These results do, however, speak to the critical issue of the timing of the onset of contextual influence. Gerrig and Bortfeld's results suggest that context may have a very early, perhaps even immediate, influence on the process of interpreting a noun-noun combination. Gagné and Spalding's results, on the other hand, suggest that if context has an influence on this interpretive process, this influence is too late to override the effect of relation frequency. It is important to note, however, that the very coarse measures these researchers have used, such as overall sentence reading time, do not reveal the detailed time-course of processing. In the present paper, I describe research using more precise measures to uncover the timing of contextual influence. To provide this greater temporal precision, I measured the eye movements of readers as they encountered critical noun-noun combinations. The eye movement record provides a detailed account of moment-to-moment processing difficulty (Rayner, 1998; Staub & Rayner, 2007), and can thereby help to determine the exact point(s) at which readers experience difficulty processing a noun-noun combination.

The onset of contextual influence is related to the onset of the semantic processing difficulty that requires the context for its resolution. Thus, the question of when context is used in the interpretation of noun-noun combinations can be decomposed into two

more specific questions. First, at what point is the difficulty of interpreting a noun-noun combination evident in the eye-tracking record? Second, at what point does potentially useful information available in the preceding context begin to ameliorate this difficulty? A strong version of the independence view described above would predict that a sufficiently supportive context might eliminate this processing difficulty altogether. Conversely, a strong version of the interdependence view would predict that even a strongly supportive context would not reduce this difficulty. However, the temporal resolution of the eye movement methodology makes it possible to assess a third possibility, namely that there is an early stage during which the process of interpretation of a noun-noun combination proceeds in isolation from context, and that context has an effect at a later stage.

Several recent studies have investigated the time-course with which semantic plausibility affects eye movements in reading (Filik, 2007; Rayner, Warren, Juhasz, & Liversedge, 2004; Staub, Rayner, Pollatsek, Hyönä, & Majewski, 2007; Warren & McConnell, 2007; Warren, McConnell, & Rayner, under review). These studies have generally compared reading times on a critical word and subsequent material when this word is plausible, given the preceding context, and when this word is implausible, as in the following pair from Rayner et al., 2004; the critical word is italicized:

1a. John used a knife to chop the large *carrots* for dinner.

1b. John used a pump to inflate the large *carrots* for dinner.

It appears that implausibility has a very rapid effect, with inflated reading times appearing as early as the reader's first fixation on the critical word in some studies (Staub et al., 2007; Warren & McConnell, 2007). In addition, there is evidence that more severe

plausibility violations have earlier effects than less severe violations (Rayner et al., 2004; Warren & McConnell, 2007). All of these studies also report continuing effects of implausibility on reading of subsequent material.

Warren et al. (under review) took a first step towards studying the role of context in the plausibility sentences similar to those used by Rayner et al. (2004). Warren et al. (under review) presented subjects with sentences such as 1b above preceded by either a real-world context (e.g., cooking during a camping trip) in which inflating carrots remains problematic or a fantasy context (e.g., inflating a Bugs Bunny balloon for a parade) in which inflating carrots would make sense. One early eye-tracking measure on the critical word (e.g., *carrots*) was longer in both the real-world and fantasy contexts. That is, the fantasy context did not facilitate early processing of the critical word. According to later measures, however, the critical word was read faster in the fantasy context. This suggests a later influence of context.

Overall, the findings from these eye movement studies suggest that a processing cost might appear as early as the first fixation on the second noun in a difficult-to-interpret noun-noun combination, compared to a condition in which the same noun appears in an easily interpretable phrase. It may then be asked whether this processing cost is reduced from the outset when a preceding context provides a possible interpretation of the critical combination, whether context does not reduce the processing cost at all, or whether context reduces processing cost only after an initial period of processing difficulty. This last hypothesis, consistent with the Warren et al. (under review) results, would predict that even in a supportive context, difficulty in processing such a combination would appear in the earliest eye movement measures (i.e., first

fixation duration and first pass time, as described below), but would also predict that such a context should reduce overall difficulty, as reflected in later eye movement measures (i.e., go-past time, total time).

It is important to note that the stimuli of Gagné and Spalding (2004) do not provide the strongest possible test of contextual influence. Combinations such as *mountain magazine*, *water bird*, *paper equipment*, and even *plastic crisis* don't require a context for their interpretation; it is clearly possible for readers to construct meaningful interpretations of these combinations without taking the context into consideration. The present study investigates the contribution of context to the interpretation of noun-noun combinations in a situation in which the potential influence of context is maximized. The influence of context can be maximized, or at least greatly increased, by using combinations with very weak interpretations in isolation, such as *olive signals* and *vacuum butter* (Gagné & Shoben, 1997). Because participants have great difficulty assigning meaning to such combinations in isolation, Gagné and Shoben referred to them as "nonsense" combinations. While these combinations are certainly difficult to interpret in isolation, strong contextual support can render these combinations meaningful, as will be demonstrated below. Thus I refer to these combinations as "near-nonsense."

As just outlined, Experiment 1 investigated the processing of such near-nonsense combinations in normal reading, using eye-tracking. But while eye movements provide a fine-grained measure of on-line processing difficulty, they do not directly answer the question of whether readers ultimately are able to make sense of these near-nonsense combinations, and if so, what meanings they assign to these combinations. Therefore, in addition to the eye-movement experiment, I conducted a questionnaire experiment that

was meant to assess more directly how these combinations were likely to be interpreted, both in and out of context. This questionnaire experiment is presented as Experiment 2 below.

CHAPTER 2

EXPERIMENT 1

Method

Participants

Forty undergraduate and graduate students at the University of Massachusetts, Amherst with normal or corrected-to-normal vision (with soft contact lenses) participated in the study. All were native speakers of American English. All were naïve concerning the purpose of the experiment; they received either course credit or a nominal sum.

Materials

I constructed 24 experimental items that varied on two factors: word type (noun-noun combinations vs. adjective-noun combinations) and context type (supportive vs. neutral). Because the noun-noun combinations were difficult to render meaningful in isolation (e.g., *dictionary treatment*), I label them as "near-nonsense". For each near-nonsense noun-noun combination, I created a corresponding, easy-to-interpret adjective-noun combination (e.g., *rough treatment*). The second word of these controls matched that of the noun-noun combination. Each of the 24 experimental items consisted of two sentences. The second sentence of each experimental item contained either a near-nonsense noun-noun combination or its associated adjective-noun combination. The first sentence provided either a context that supported the understanding of the noun-noun combination or a neutral context. The four conditions for an experimental item are illustrated in (2a-d), below.

2a. When he gets angry with me he uses profanity that I have to look up. He gave me the dictionary treatment just last night. (noun-noun combination, supportive context)

2b. He has given me several things over the course of the past few weeks. He gave me the dictionary treatment just last night. (noun-noun combination, neutral context)

2c. When he gets angry with me he uses profanity that I have to look up. He gave me the rough treatment just last night. (adjective-noun combination, supportive context)

2d. He has given me several things over the course of the past few weeks. He gave me the rough treatment just last night. (adjective-noun combination, neutral context)

The supportive context sentence and the neutral context sentence were roughly matched on length; on average, the neutral context sentence was less than two characters longer than the supportive context sentence (means of 69.5 and 67.9 characters). As the study was not designed to investigate lexical priming, I constructed the supportive and neutral contexts so that no word lexically primed either of the words of the noun-noun combination. The norms of Nelson, McEvoy, & Schreiber (1998) and the Edinburgh Associative Thesaurus (Kiss, Armstrong, Milroy, & Piper, 1973) were used to assess lexical association. In these sources, the critical words never appeared in association with any of the words in the context sentences.

The mean length of the initial nouns in the noun-noun combinations was 5.58 characters, while the corresponding adjectives averaged 6.17 characters. All within-item differences were within 4 characters with two exceptions: a 5-character difference for

dictionary/rough treatment and an 8-character difference for *curry/international electronics*. All experimental materials are included in the Appendix.

I combined the 24 experimental items with 48 filler items, which were items from unrelated experiments that did not investigate the processing of noun-noun combinations. Twenty-four of the fillers contained a reduced relative clause garden path construction (Ferreira & Clifton, 1986; Frazier & Rayner, 1982), which was anticipated to be difficult to understand and would help draw attention away from any oddity of the experimental items. To check for understanding, participants answered a yes/no comprehension question on a third of the items (half of which required a "yes" response). The comprehension questions asked about content from all parts of experimental sentences so that participants could not focus exclusively on the noun-noun combinations.

Procedure and Apparatus

Eye movements were recorded on an SR Research Eyelink 1000 tracker. The sampling rate was 1000 Hz. Only the right eye's movement was monitored, although viewing was binocular. All sentences were displayed on a maximum of two lines with a maximum line length of 101 characters. Stimuli were displayed on a 17-inch Viewsonic monitor. Participants were seated 60 cm from the computer screen. At this distance, 3.2 characters subtend 1° of visual angle, with the eye-tracker having an angular resolution of 10-30 min of arc. The experiment took approximately 30 minutes.

Participants were asked to read for comprehension and were told that after some of the items they would be required to respond to a comprehension question by pressing a *yes* or *no* button. They were also told that some of the items might be "a little weird." Once the participant was seated at the eye-tracker, the tracker was aligned and calibrated.

Results

The analysis of eye movement data focused on three regions in the second sentence of each experimental item. The primary region of interest was the *head noun* of the combination (e.g., *treatment* in *dictionary treatment*). If the interdependence view is correct about the influence of context, there should be no effect of context type on processing of this noun. On the other hand, the independence view predicts that a supportive context should reduce processing difficulty. To investigate ongoing processing difficulty caused by the experimental manipulations, I also examined reading of a *spillover* region consisting of the remainder of the sentence following the head noun (e.g., *just last night).* To ensure that there were no unintended effects of my experimental manipulations prior to the head noun of the combination, I examined reading time on the *modifier* of the combination (e.g., *dictionary* in *dictionary treatment*, *rough* in *rough treatment*).

The following reading measures were computed: *first fixation duration*, *first pass time* (referred to as *gaze duration* for single-word regions), *go-past time*, *total time*, and *percent regressions out*. First fixation duration is the duration of the first fixation in a region, regardless of whether it is the only fixation in the region or the first of several fixations. For single-word regions, this measure indexes lexical processing difficulty (e.g., Reichle, Rayner, & Pollatsek, 2003). First pass time/gaze duration is the sum of all fixations in a region prior to leaving the region for the first time in either direction. Go-past time starts summing fixations starting with the first fixation in the region until the region is exited to the right. This sum includes fixations to the left of the region after a regressive eye movement plus any fixations that then reenter the region before moving on

to the right. Total time is the sum of all fixations in a region, which can include regressions back into the region from text further along in the item. Percent regressions out gives the percent of trials on which readers made a regressive eye movement out of a region on the first pass through the region.

Four of the measures rest on a continuum ranging from early processing to late processing. First fixation duration represents the first direct contact with the region and corresponds to the earliest processing that takes place on that region. (Though there is clear evidence of parafoveal processing of a word while the eyes are fixating the previous word, it appears that this processing does not include semantic or syntactic processing; see, e.g., Staub & Rayner, 2007). Because first pass time/gaze duration includes subsequent fixations on the same word or region, slower-acting influences on processing difficulty may be able to exert an influence on this measure. Similarly, because go-past time includes fixations made after a regressive eye movement from the region of interest, its duration may be taken to reflect even slower-acting influences. As an example, it is sometimes the case that the effect of syntactic garden-pathing appears in regression-based measures such as go-past time, but not in gaze duration/first pass time (Clifton, Staub, & Rayner, 2007). Finally, because total time includes regressions into the region from subsequent text, it may be considered an even later measure of processing.

Comprehension was high. Participants averaged 89% accuracy on the comprehension questions. Trials with track losses were eliminated (less than 5% of trials) before analysis. Short fixations (< 80 ms) located within one character of an adjacent fixation (previous or subsequent) were assimilated into the adjacent fixation. The same procedure was used for extremely short fixations (< 40 ms) within three characters of a

neighboring fixation. Any remaining fixations less than 80 ms were excluded as well as fixations longer than 800 ms. Useful information is most likely not gleaned from fixations less than 80 ms while extremely long fixations (> 800 ms) likely reflect track losses (see Rayner & Pollatsek, 1989). These criteria resulted in the elimination of less than 2.5% of all fixations.

For each measure on each region, I performed two ANOVAs, treating participants (F1) and items (F2) as random effects variables. If an effect is reported as not significant, the values of both F tests were less than 1.5. In both the participants and items analyses, word type (noun or adjective as the modifier of the combination: noun-noun or adjective-noun) and context (supportive or neutral) were treated as within factors. Table 1 shows the means for each of the dependent measures in each condition, on each region. For consistency of presentation, the response times and percent regressions are presented in the following order: for word type, noun-nouns are given prior to the adjective-nouns; for context type, the supportive context is given prior to the neutral context.

Modifier Region

Because the modifying noun in a noun-noun combination was not matched on length or frequency with the modifying adjective in an adjective-noun combination, the modifying noun and adjective cannot be compared directly. For this reason, I do not report main effects of word type (noun vs. adjective) in the modifier region.

Context type significantly affected the duration of the first fixation on the modifier, with shorter fixations in the supportive context (218.5 ms vs. 232.5 ms): $F1(1, 39) = 6.67$, $p = .01$; $F2(1, 23) = 6.32$, $p < .05$. However, no effects of context approached

significance for gaze duration or go-past times, and indeed, go-past time was numerically shorter in the neutral context than in the supportive context. Therefore, I suspect that the first fixation result is specious, perhaps arising from slight differences in the initial landing position of the eyes. There were also no significant interaction effects on the first fixation, gaze, or go-past measures.

On the total time measure, there was a significant main effect of context, with much shorter times in the supportive context (383 ms vs. 446 ms: $F1(1, 39) = 14.87$, $p < .001$; $F2(1, 23) = 10.01$, $p < .001$). There was also a significant effect of the interaction between word type and context on total time ($F1(1, 39) = 17.63$, $p < .001$; $F2(1, 23) = 9.45$, $p < .01$). These effects can be explained by the differential probabilities of regressing back into the region from later in the sentence, as discussed below. There were no significant effects on the regressions out measure.

Noun Region

There were significant main effects of word type on all four reading time measures, with shorter reading times for the critical noun in the adjective-noun condition: first fixation (247 ms vs. 223.5 ms: $F1(1, 39) = 23.67$, $p < .001$; $F2(1, 23) = 12.57$, $p < .01$); gaze duration (305.5 ms vs. 262 ms: $F1(1, 39) = 23.22$, $p < .001$; $F2(1, 23) = 20.80$, $p < .001$); go-past time (506 ms vs. 363.5 ms: $F1(1, 39) = 30.92$, $p < .001$; $F2(1, 23) = 16.62$, $p < .001$); and total time (485 ms vs. 331 ms: $F1(1, 39) = 84.66$, $p < .001$; $F2(1, 23) = 53.83$, $p < .001$).

The effect of context on early measures (first fixation and gaze) was not significant. The effect of context on go-past time was marginally significant, and context

had a fully significant effect on total time, with shorter reading times in the supportive context: go-past (409 ms vs. 461 ms: $F1(1, 39) = 3.66$, $p = .06$; $F2(1, 23) = 3.18$, $p = .09$); total time (387 ms vs. 429 ms: $F1(1, 39) = 10.90$, $p < .01$; $F2(1, 23) = 6.39$, $p < .05$).

The interaction of word type and context did not significantly affect the early measures. However, interaction effects on go-past and total time were significant: go past ($F1(1, 39) = 6.56$, $p < .05$; $F2(1, 23) = 5.66$, $p < .05$); total time ($F1(1,39) = 5.92$, $p < .05$; $F2(1, 23) = 9.47$, $p < .01$). For example, on the go-past measure the supportive context resulted in slightly longer reading times for the noun in the adjective-noun condition (367 ms vs. 360 ms), while the supportive context dramatically reduced reading time for the noun in the noun-noun condition (451 ms vs. 561 ms). The simple effect for the adjective-noun condition was not significant while the simple effect for the noun-noun condition was significant ($F1(1, 39) = 11.68$, $p < .01$; $F2(1, 23) = 9.90$, $p < .01$).

There was a significant main effect of word type on regressions out, with fewer regressions out of the noun region in the adjective-noun condition (33% vs. 22%: $F1(1, 39) = 14.90$, $p < .001$; $F2(1, 23) = 6.54$, $p < .05$). There was also a significant main effect of context with fewer regressions out of the noun when it occurred in the supportive context (22% vs. 32%: $F1(1, 39) = 8.49$, $p < .01$; $F2(1, 23) = 4.50$, $p = .05$). The interaction between these factors was not significant.

Spillover Region

There were no significant effects of the experimental manipulations on the early measures of processing difficulty on the spillover region; there was only a slight hint of an interaction effect on first fixation duration: $F1(1, 39) = 2.41$, $p = .13$; $F2(1, 23) = 1.91$,

p = .18. Go-past time, however, was significantly affected by word type, with shorter times in the adjective-noun conditions (1452 ms vs. 1172 ms: $F1(1, 39) = 11.27$, $p < .05$; $F2(1, 23) = 12.21$, $p < .05$), and was marginally affected by context, with shorter times in the supportive context (1255 ms vs. 1369 ms: $F1(1, 39) = 3.19$, $p = .08$; $F2(1, 23) = 2.75$, $p = .11$). Word type also significantly affected total time, with shorter times in the adjective-noun conditions (865 ms vs. 765 ms: $F1(1, 39) = 8.79$, $p < .01$; $F2(1, 23) = 7.98$, $p = .01$), though context did not significantly affect total time. There were no interaction effects on go-past time or total time. Finally, there were also fewer regressions out of the spillover region in the adjective-noun conditions (50% vs. 35%: $F1(1, 39) = 22.87$, $p < .001$; $F2(1, 23) = 15.66$, $p = .001$). There was also a marginal effect of context type with fewer regressions out of the region in the supportive context (39% vs. 46%: $F1(1, 39) = 3.97$, $p = .053$; $F2(1, 23) = 3.64$, $p = .07$), and a marginal interaction of context and word type ($F1(1, 39) = 3.33$, $p = .08$; $F2(1, 23) = 7.36$, $p = .01$). The simple effect for adjective-nouns was not significant while for noun-nouns it was significant ($F1(1, 39) = 6.95$, $p < .05$; $F2(1, 23) = 8.19$, $p < .01$).

Combined Region: Noun and Spillover

As noted above, there was a marginal effect of the word type x context interaction on regressions out of the spillover region. There was also a nonsignificant trend indicative of the same pattern on the noun region ($F1(1, 39) = 2.47$, $p = .12$; $F2(1, 23) = 0.82$, $p = .37$). In order to assess whether these trends mask a genuine experimental effect, I re-analyzed the regressions data after creating a combined region comprised of these two regions (e.g., *treatment just last night.*). In the four conditions, the means on

the percent regressions measure were as follows: neutral noun-noun 69%; supportive noun-noun 49%; neutral adjective-noun 41%; supportive adjective-noun 40%. The interaction effect was now highly significant ($F1(1, 39) = 6.83$, $p = .01$; $F2(1, 23) = 9.81$, $p < .01$).

Second Sentence as a Single Region

In order to compare these results as directly as possible with those of Gerrig and Bortfeld (1999), who computed whole-sentence reading times for the sentences containing their critical noun-noun combinations, I also computed reading times for the second sentence as a whole. (The measure used by Gagné and Spalding (2004) was the time to respond to a comprehension question, which does not have an analogue in the eye-movement record.) Because Gerrig and Bortfeld had participants read one sentence at a time, using a key press to proceed to the next sentence, their participants were unable to regress back to reread previous sentences. In my two-sentence stimuli, participants could regress back into the first sentence. In order to compare my data to Gerrig and Bortfeld's measure, I used two different measures. As a first approximation I measured go-past time for the second sentence as a whole. This measure includes all the trials but allows regressions from the second sentence back into the first sentence, so I also measured the first pass time on the second sentence for those trials in which readers did not regress back to the first sentence. For the go-past measure on the second sentence, there were significant effects of both word type and context. The adjective-noun condition was read faster than the noun-noun condition (2659 ms vs. 2223 ms: $F1(1, 39) = 22.76$, $p < .001$; $F2(1, 23) = 24.15$, $p < .001$), and the second sentence was read more quickly when it

followed the supportive context sentence (2343 ms vs. 2539 ms: $F1(1, 39) = 10.01$, $p < .01$; $F2(1, 23) = 5.60$, $p < .05$). The interaction was in the expected direction (i.e., larger effects of context for the noun-noun sentences than for the adjective-noun sentences), but did not approach significance ($F1(1, 39) = 1.634$, $p = .21$; $F2(1, 23) = 1.53$, $p = .23$). For the first pass measure excluding trials with regressions, there was an effect of word type, with faster reading in the adjective-noun conditions (2301 ms vs. 1823 ms: $F1(1, 38) = 71.07$, $p < .001$; $F2(1, 23) = 38.45$, $p < .001$), a marginal effect of context, with faster reading in the supportive context (2016 ms vs. 2108 ms: $F1(1, 38) = 3.79$, $p = .06$; $F2(1, 23) = 6.13$, $p < .05$), and again only a hint of an interaction effect ($F1(1, 38) = 1.90$, $p = .18$; $F2(1, 23) = 1.47$, $p = .24$).

Discussion

The pattern of results from this experiment is quite clear. First, the critical noun was indeed more difficult to process when it appeared in a near-nonsense noun-noun combination than when it appeared in an adjective-noun combination. This processing difficulty was apparent in the form of inflated reading times on the noun itself, beginning with the first fixation on the word but also appearing in gaze duration, go-past time, and total time, and in the form of inflated reading times on the subsequent material, appearing in the later eye movement measures.

Second, it is clear that the supportive context reduced this processing difficulty: significant interactions between word type and context appeared on the go-past and total time measures on the critical noun, and on the percent regressions measure for a region consisting of the noun plus all subsequent material. The latter result is especially telling: when the preceding context did not support a specific interpretation of the critical

compound, readers executed a regressive eye movement out of this region 69% of the time in the noun-noun condition. But when the preceding context did support such an interpretation, they executed a regressive eye movement only 49% of the time (compared to 41% and 40% in the adjective-noun conditions). Importantly, this modulation of processing difficulty first appeared in regression-based measures on the critical noun: there was essentially no hint of an interaction with context on the first fixation or gaze duration measures. As expected, there was no compelling evidence that context type had any effect on the processing of the adjective-noun combinations.

These results clearly support a two-stage model of the processing of noun-noun combinations. The early eye-tracking measures show that the supportive context did not alleviate the initial processing difficulty for the near-nonsense noun-noun combinations. As indexed by the later eye-tracking measures, however, the supportive context did alleviate the subsequent processing difficulty for these same combinations. Apparently, there is an early stage of interpretation that proceeds in isolation from context, and a later stage in which context is able to ameliorate processing difficulty. Thus, the results are inconsistent with simple versions of either the independence (Gerrig & Bortfeld, 1999) or interdependence (Gagné & Spalding, 2004) views.

Recall that much of the previous research in this area relied on an overall reading time measure. As a point of comparison with that research, the two total sentence reading time measures for the current data seem to provide limited support for the independence view: The context manipulation was able to modulate the difficulty associated with the noun-noun combinations only nonsignificantly. But an examination

of the detailed time-course of processing makes clear that this finding masks a more complex process.

CHAPTER 3

EXPERIMENT 2

The primary goal of Experiment 2 was to provide a direct measure of participants' understanding of my near-nonsense combinations. After reading these noun-noun combinations in various contexts, participants were asked to write out the meanings for the near-nonsense combinations used in Experiment 1. Participants were also asked to rate the difficulty of determining these meanings. Comparison of these ratings would indicate whether the supportive context results in a subjective change in the difficulty of the process of interpretation.

A secondary goal was to measure whether participants can make sense of the near-nonsense combinations in isolation. Participants in Gagné & Shoben's (1997) study judged two of my near-nonsense combinations (e. g., *olive signals* and *vacuum butter*) as nonsense when exposed to them for approximately 1200 ms. I was interested in whether giving participants more exposure time, approximately 35 seconds, would enable them to make sense of these demanding combinations in isolation.

Method

Participants

Thirty-three undergraduates at the University of Massachusetts Amherst participated in the study. All were native English speakers and received course credit for their participation.

Materials

The near-nonsense combinations used in Experiment 1 appeared in written form in one of three contexts. The supportive and neutral contexts were identical to those from Experiment 1. A new context, called an isolated context, provided no information that could assist in comprehending the critical combination. In this way, it approximated the situation of encountering a near-nonsense combination in isolation. For example, after reading one of the contexts below, participants were asked the question, "What is the first meaning you think of for 'dictionary treatment'?"

3a. When he gets angry with me he uses profanity that I have to look up. He gave me the dictionary treatment just last night. (supportive context)

3b. He has given me several things over the course of the past few weeks. He gave me the dictionary treatment just last night. (neutral context)

3c. In the fashion world, there are secret rules for the designers to follow. Meeting the standards of clothespin integrity is difficult. (isolated context)

The isolated context (3c) conveys information relevant to the combination *clothespin integrity* instead of *dictionary treatment*, providing no information relevant to the meaning of *dictionary treatment*.

Thirty-six filler stimuli were included in order to prevent participants from focusing exclusively on the noun-noun combinations as they read. The comprehension questions that were paired with these fillers asked about different parts of the sentences.

For example, the question paired with the sentence below was "What type of surface did the athlete trip on?"

4. The athlete tripped and fell when one of the metal protrusions on the bottom of his shoe caught on the artificial turf. He had a cleat hiccup in front of everyone.

Procedure

Participants were tested in groups in sessions lasting approximately 50 minutes. They were given a booklet containing 128 pages, and were instructed that for every pair of pages in the booklet, the first page contained a two-sentence passage and the second page contained a question and a rating scale for the difficulty of the question (1=easy, 7=difficult). Each booklet consisted of four practice stimuli followed by a pseudo-randomized ordering of the 24 experimental items and 36 filler items. The experimental items were arranged into three counterbalanced lists so that each participant completed eight items in each of the three conditions, and each item was rated by between 10-12 participants in each of its three conditions. Participants were instructed that occasionally the question they were to answer had no relation to the two sentences they just read. In this case, the participants were instructed to answer the question as well as possible.

Participants were given 35 seconds to read a two-sentence stimulus, turn the page, write their answer, and rate the difficulty. Once they turned to the question page, they were not allowed to turn back to re-read the two-sentence stimulus. An experimenter kept time and told participants when the 35 seconds were up and when they could turn to the next two-sentence stimulus. If a participant finished a question before the 35 seconds were up, they were to wait until the experimenter instructed them to turn to the next stimulus.

Results

As in Experiment 1, I performed two ANOVAs, treating participants (F1) and items (F2) as random effects variables. If both F tests were less than 1.5, then the effect was reported as not significant. In both the participants and items analyses, context (supportive, neutral, or isolated) was treated as a within factor.

Two judges independently classified the participants' written meanings into four categories: no answer; unintelligible answer ("something about a dictionary" for *dictionary treatment*); target meaning, the meaning based on the supportive context ("receiving a flurry of unknown curse words"); and non-target meaning, another meaning not based on the supportive context ("being corrected in vocabulary and grammar"). Any disagreements between the judges (less than 2% of all judgments) were resolved through discussion.

The percentage of responses in each of the four categories, for each of the three contexts, is given in Table 2. Across the three conditions there was a significant difference in the proportion of target meanings, $F1(2, 64) = 154.79$, $p < .001$; $F2(2, 46) = 100.30$, $p < .001$. All pairwise tests were also significant: supportive-neutral, $F1(1, 32) = 148.92$, $p < .001$; $F2(1, 23) = 90.49$, $p < .001$: supportive-isolated, $F1(1, 32) = 377.47$, $p < .001$; $F2(1, 23) = 181.09$, $p < .001$: and neutral-isolated, $F1(1, 32) = 19.69$, $p < .001$; $F2(1, 23) = 16.19$, $p = .001$.

I also computed the percentage of trials on which any coherent meaning, i.e., the sum of the target and non-target meanings, was produced for each of the three contexts:

supportive, 94.7; neutral, 83.7; and isolated, 76.9. There was a significant difference across the three contexts, $F1(2, 64) = 12.22$, $p < .001$; $F2(2, 46) = 11.48$, p , .001. The supportive context was significantly different than either of the other contexts: supportive-neutral, $F1(1, 32) = 11.73$, $p < .05$; $F2(1, 23) = 8.80$, $p < .01$: supportive-isolated, $F1(1, 32) = 23.27$, $p < .001$; $F2(1, 23) = 22.02$, $p < .001$. The difference between the neutral and isolated contexts did not approach significance, $F1(1, 32) = 2.96$, $p = .10$; $F2(1, 23) = 2.93$, $p = .10$.

A second analysis determined whether the type of context modified the subjective difficulty of interpretation. The mean difficulty ratings for the three contexts were as follows: supportive, 2.6; neutral, 4.1; isolated, 4.7. There was a significant difference in difficulty rating among the contexts, $F1(2, 64) = 67.71$, $p < .001$; $F2(2, 46) = 148.78$, $p < .001$. All pairwise tests were significant: supportive-neutral, $F1(1, 32) = 90.29$, $p < .001$; $F2(1, 23) = 142.54$, $p < .001$: supportive-isolated, $F1(1, 32) = 93.34$, $p < .001$; $F2(1, 23) = 315.89$, $p < .001$: and neutral-isolated, $F1(1, 32) = 9.91$, $p < .05$; $F2(1, 23) = 18.15$, $p < .001$.

Discussion

The results of Experiment 2 support the intuition that the information contained in the supportive contexts used in Experiment 1 facilitates interpretation of the noun-noun combinations. First, participants reported that generating a meaning was easier when given a supportive context. Second, participants were very likely to arrive at the target meaning with a supportive context, doing so on 79% of trials (compared to 33% and 15% in the neutral and isolated conditions, respectively).

It is interesting to note that a neutral context led to a target meaning more often than an isolated context, and the meaning-generation process was also judged to be easier. This result suggests one of two possibilities. First, although the results show that a neutral context does indeed carry less useful information than a supportive context, the neutral context might still be helpful in the process of meaning creation. For example, the neutral context for the combination *olive signals* (*I walked around the party for quite a while before I noticed it. I was receiving olive signals from across the room.*) still provides vague information that is generally consistent with the target meaning (*flirting over a martini*). The context suggests a party situation where flirting might occur and alcohol might be served. This example is extreme; other neutral contexts provide much less useful information. Note that although the neutral context might carry helpful information, for the design of Experiment 1 it was critical only that the supportive context (*The woman with the martini was looking at me with interest. I was receiving olive signals from across the room.*) provide more support than the neutral context.

Second, the isolated context might be disruptive. The isolated context for *olive signals* (*In the fashion world, there are secret rules for the designers to follow. Meeting the standards of clothespin integrity is difficult.*) does not mention the relevant combination. There might be a cost involved in disregarding the isolated context when creating a meaning for *olive signals*. In this regard, the isolated context differs from the isolated conditions of previous researchers. This experiment was not designed to distinguish between these two possibilities.

Note that, although meaning creation in an isolated context was judged to be more difficult than meaning creation in a neutral context, participants were still able to produce

a coherent meaning on a surprising 77% of the trials. This result contrasts with the conclusions of Gagné and Shoben (1997). Participants in their experiments quickly judged whether noun-noun combinations made sense. Under these conditions, their participants deemed combinations such as *vacuum butter* and *olive signals* to be nonsense; whereas, my participants generated quite creative meanings for such combinations. For example, participants gave the following meanings for *vacuum butter*: "[butter] so good it kept being eaten" (i.e., disappearing); "the oily substance between the parts of a vacuum"; and "something easy to vacuum".[2]

The amount of time participants were given to generate a meaning differed substantially between my study and that of Gagné and Shoben (1997). The participants in Gagné and Shoben (1997) took an average of approximately 1097 ms and 1184 ms in Experiments 1 and 3, respectively, to decide whether a combination was sense or nonsense. My participants, in contrast, had up to 35 seconds to read the context sentences and create and write a meaning. It is hard to directly compare these times, as the Gagné and Shoben (1997) participants did not have to explicitly construct and communicate a meaning for the combinations, nor did they have to read two-sentence passages. The results do indicate, however, that participants can create coherent meanings for these difficult combinations if they are given enough time, even the modest amount afforded in my study.[3]

CHAPTER 5

GENERAL DISCUSSION

Two experiments investigated the questions of whether and when contextual information is used in the interpretation of noun-noun combinations. In Experiment 1, participants' eye movements were monitored as they read two-sentence passages containing near-nonsense noun-noun combinations. These combinations proved more difficult to process than adjective-noun combinations in which the second word was identical, as indicated by inflated reading time on the critical noun and subsequent material. More interestingly, a preceding context designed to encourage a particular interpretation of the critical compound did not modulate this increase in reading time in the earliest eye movement measures, i.e., first fixation duration and gaze duration on the critical noun, but did modulate the increase in reading time in later eye movement measures, i.e., go-past time and total time on the critical noun, and also modulated the frequency of regressive eye movements. Experiment 2 validated the materials used in Experiment 1, showing that the supportive contexts made it more likely that participants would reach a specified "target" interpretation, made it more likely that they would reach any interpretation at all, and made the task of reaching an interpretation subjectively easier. Thus, the effect of context on reading time in Experiment 1 can indeed be interpreted as reflecting a reduction in the difficulty of meaning-creation.

The finding that the context has a delayed influence on the interpretation of noun-noun combinations is consistent with the fantasy context study of Warren et al. (under review). It is also consistent with the conception of semantic interpretation in sentence processing as a *compositional* process (see, e.g., Pylkkänen & McElree, 2007).

Compositionality holds, among other things, that more local semantic computations within a sentence are performed before more global ones, where the relevant notion of locality is defined with respect to syntactic structure. Thus, a particular noun phrase involving an adjective-noun or noun-noun combination would receive an independent interpretation, which would then be fed into a computation of the meaning of a larger unit such as a verb phrase, and so on. Finally, the comprehender may bring information from the discourse to bear on the interpretation of a compositionally-derived sentence meaning.

However, it is worth noting that in a previous eye movement experiment, Filik (2007) found that a sufficiently well-developed preceding context can completely eliminate the processing difficulty associated with implausibility. In Filik's experiment, readers encountered a noun that served as an implausible theme for the preceding verb, or an implausible theme given the combination of subject and verb (e.g., *The mouse picked up the dynamite*…). The preceding sentence either did nothing to explain this apparent implausibility, or, in the critical conditions, it established a licensing fictional context (e.g., one that referred to the cartoon cat and mouse Tom and Jerry). The key finding was that the effect of implausibility in the eye movement record was totally eliminated with the licensing context.

This result appears to contradict my finding of a delayed influence of context, and it would be tempting to interpret the difference in results in terms of a difference between noun-noun combinations on the one hand and more global implausibilities on the other. However, I think that given the details of Filik's (2007) data, this conclusion is not warranted. In the Filik study, there were no significant effects of implausibility on the

critical noun itself, which stands in sharp contrast both to the present study and to other eye movement studies of the processing of implausibility (Rayner et al., 2004; Staub et al., 2007; Warren & McConnell, 2007). All effects of the plausibility manipulation emerged further downstream. Moreover, there were no significant effects on the frequency of regressive eye movements, even from downstream regions, which, once again, stands in contrast to the present study and to previous studies. In short, the plausibility manipulation in Filik's study only affected first-pass reading time measures on the post-critical region. This suggests that the plausibility manipulation was not very strong, which may explain why contextual override was possible.

Interestingly, the eye movement patterns upon encountering the second noun in a near-nonsense noun-noun combination, in the absence of a supportive context, look essentially like the patterns found in previous experiments that manipulated the plausibility of a noun in context. As noted above, these studies have varied the severity of the implausibility, in general finding first fixation effects only with the most severe implausibilities. The present study did obtain a very sizable first fixation effect. On the compositional perspective outlined above, this finding has a clear explanation, namely that the processing of the noun phrase itself takes place independently in a first stage of processing that is followed by a second stage that incorporates the computed meaning of the noun phrase into the more global aspects of the sentence.

Implications for a Theory of Noun-Noun Combination

I concur with relation-linking theory in assuming that the initial stage of interpreting a noun-noun combination involves identifying a potential connecting relation

between the two concepts (e.g., producing the relation *uses* in "a treatment *using* a dictionary"). I suggest, however, that context must be used to create a more specific and detailed way for the two concepts to interact. The contextual information shows how the general connecting relation between the two concepts can be implemented in an actual situation (e.g., "treatment using unknown curse words that require the use of a dictionary to determine their meaning"). Embedding the relation in a context may proceed in two stages. In the first stage, the comprehender attempts to fit the relation into the most easily accessible context, i.e., general knowledge and well-established discourse information. If this attempt fails, the comprehender might then attempt to construct a context in which the relation can be realized.

The present eye-movement results appear to map onto this two-stage model in a relatively transparent way: The early measures such as first fixation and gaze duration reflect difficulty in determining a relation and checking it against the default context, while the later measures map onto difficulty constructing a novel context. A compound such as *dictionary treatment* usually fails to find a specific interpretation in the default context, giving rise to difficulty in the earliest eye movement measures regardless of the information presented in the preceding sentence. The information in the preceding sentence may ease the task of context construction in the second stage. Finally, I suspect that the process may restart if no suitable context can be constructed for the currently selected relation. This theory maintains the core ideas of relation-linking theory, but at the same time provides an explicit role for context in the creation of meaning for noun-noun combinations.

Table 1. Eye-tracking measures from Experiment 1.

Condition	Region		
	Modifier	Noun	Spillover
	First Fixation (ms)		
Neutral noun-noun	234	250	250
Supportive noun-noun	219	244	235
Neutral adj-noun	231	225	233
Supportive adj-noun	218	223	236
	First-Pass (ms)		
Neutral noun-noun	277	307	622
Supportive noun-noun	260	304	601
Neutral adj-noun	258	263	664
Supportive adj-noun	254	261	616
	Go-Past (ms)		
Neutral noun-noun	381	561	1534
Supportive noun-noun	387	451	1370
Neutral adj-noun	315	360	1203
Supportive adj-noun	343	367	1140
	Total Time (ms)		
Neutral noun-noun	529	522	871
Supportive noun-noun	405	447	859
Neutral adj-noun	363	336	782
Supportive adj-noun	361	326	747
	Regressions Out (percent)		
Neutral noun-noun	24	39	56
Supportive noun-noun	26	26	43
Neutral adj-noun	20	25	35
Supportive adj-noun	21	18	34

Table 2. Percent of responses in each category in Experiment 2, by context type.

	Supportive	Neutral	Isolated
No Answer	0	2.3	3.4
Unintelligible Answer	5.7	14.0	19.3
Target Meaning	79.2	33.3	14.8
Non-target Meaning	15.5	50.4	62.1

Notes

[1] Throughout this paper, I follow the common practice in the conceptual combination literature and assume that nouns and adjectives correspond to concepts.

[2] *Vacuum butter* and *olive signals* were the only two combinations that overlapped with the Gagné and Shoben (1997) stimuli. Participants were able to generate a meaning for *vacuum butter* 6 of the 12 times it was encountered in the isolated context (all were non-target meanings). Participants were able to generate a meaning for *olive signals* 8 of the 10 times it was encountered in the isolated context (again, all were non-target meanings).

[3] This leads to the question of whether there are any noun-noun combinations that are truly nonsense, in the sense that they cannot be given intuitively plausibly interpretations by most native speakers. I created a random noun-noun combination generator that draws nouns from Princeton's WordNet database (http://wordnet.princeton.edu) of approximately 57,000 nouns. As an informal experiment, I created 500 random combinations and tried creating meanings for each of them by creating some relevant context. In all cases, I was ultimately successful. Although this test is inconclusive regarding the existence of a truly nonsensical combination, it leads me to believe that the actual frequency is quite low.

APPENDIX

EXPERIMENTAL ITEMS

The twenty-four experimental items for the eye-tracking study are listed below. Each sentence pair consisted of a (a) supportive or (b) neutral context sentence followed by a (c) sentence which included one of the word types separated by a slash (noun/adjective).

a. The woman with the martini was looking at me with interest. *(supportive context)*
b. I walked around the party for quite a while before I noticed it. *(neutral context)*
c. I was receiving olive/frantic signals from across the room. *(noun/adjective)*

a. The car company searched for someone to do research on novel fuels.
b. They searched all around the country to find the right one for the task.
c. They brought in an ethanol/new hire as soon as they could.

a. Being just a helper I received the least amount when the money from the herd was divided up.
b. Just like every other day of my life, I worked hard that day and received something in return.
c. I received the cowhand/standard share at the end of the day.

a. Sitting out Saturday evening, my grandmother witnessed a robbery.
b. Yesterday she spent time in all the areas of her home both inside and out.
c. She has been experiencing porch/emotional distress since then.

a. It was a new food that helped her get rid of fat.
b. There were many new foods that she tasted at the party.
c. She tried the vacuum/latest butter and liked it.

a. Everyone needs something to give them hope whey they are gloomy.
b. They do not agree on much but they agreed on this one point.
c. Everyone needs a dungeon/radiant sun in their life.

a. He was so thin that he could blow away.
b. He had a great personality but an unusual look.
c. He had a kite/lean body that was striking.

a. Whenever she had difficulty with her roommate she climbed up to a high place to think things through.
b. When we could not find her, we now knew where to look because she always went to the same place.
c. She could be found on the issue/steep rooftop during these times.

a. What he said sounded nice but was really an insult.

b. At the end of the day he said something to his buddy.
c. He gave a floor/crude compliment to his friend.

a. She rewarded him for making more of everyone's reports.
b. She has given out many things to many people for various reasons.
c. She gave him the copier/big trophy before he left for the day.

a. Whenever my father flagged the page of his novel and closed it I knew he had something serious to discuss.
b. I was out all evening with some friends having a lot of fun but I got home way beyond the time that I told my father.
c. He gave me a bookmark/long talk when I got home that evening.

a. She has been giving him subtle hints that she wants him to propose.
b. They just keep coming to him regularly whether he wants them to or not.
c. He has been receiving diamond/important messages almost daily.

a. In college I did a lot of cocaine with the guys who lived in my dorm.
b. I was close to many people in college and I divided them into named groups.
c. I still call my snort/brave buddies on a regular basis.

a. Asian countries have been educating millions of programmers.
b. Technology has gone through so many phases and changes recently.
c. We are entering the age of curry/international electronics according to many experts.

a. The activist supported the rights of the migrant farmers.
b. He has worked for the organization for many years.
c. He has engaged in salad/state politics most of his life.

a. He believes he can find happiness through mind-altering substances.
b. One thing alone seems to get him through extremely difficult times.
c. His syringe/strong hope sustains him.

a. He believes that everything his fatherland does is good.
b. He has become fixated on something to an extreme.
c. His flag/deep trance is severe.

a. When he gets angry with me he uses profanity that I have to look up.
b. He has given me several things over the course of the past few weeks.
c. He gave me the dictionary/rough treatment just last night.

a. My brother won't listen to guys but he'll do anything that a woman asks him.
b. He is a very gullible person and to this day he still has not learned his lesson.
c. He has been vulnerable to skirt/foolish persuasion all his life.

a. He thinks that he never makes a mistake so he doesn't need an eraser.

b. He has some extreme characteristics that no one knows the origin of.
c. He possesses an ink/firm confidence that is excessive.

a. She should be speaking up for herself but she remains silent.
b. She has been acting strangely and everyone has noticed it.
c. She has a tape/swollen mouth today and I don't know why.

a. The smallest disagreements between them seem to shake their devotion to each other.
b. When they first met no one could have predicted that things would turn out this way.
c. Their jello/fragile marriage is very interesting to observe.

a. He gets up late so it has slipped his mind that some people actually eat in the morning.
b. He has changed in many ways over the years, some of them good and others not so good.
c. He developed cereal/selective amnesia after he went to college.

a. He has a tough exterior but is very sensitive on the inside.
b. I didn't know what to expect when I first met him.
c. His egg/unique personality can be surprising.

REFERENCES

Downing, P. (1977). On the creation and use of English compound nouns. *Language,* 53*,* 810-842.

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

Filik, R. (in press). Contextual override of pragmatic anomalies: Evidence from eye movements. *Cognition*.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14*, 178-210.

Gagné, C. L., & Shoben, E. J. (1997). Influence of thematic relations on the comprehension of modifier-noun combinations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 71-87.

Gagné, C. L., & Spalding, T. L. (2004). Effect of discourse context and modifier relation frequency on conceptual combination. *Journal of Memory and Language*, *50*, 444-455.

Gerrig, R. J., & Bortfeld, H. (1999). Sense creation in and out of discourse contexts. *Journal of Memory and Language*, *41*, 457-468.

Kiss, G.R., Armstrong, C., Milroy, R., and Piper, J. (1973) An associative thesaurus of English and its computer analysis. In Aitken, A.J., Bailey, R.W. & Hamilton-Smith, N. (Eds.) *The Computer and Literary Studies*. Edinburgh: University Press. http://www.eat.rl.ac.uk/

Levi, J. (1978). *The syntax and semantics of complex nominals.* New York: Academic Press.

Murphy, G. L. (2004). *The big book of concepts.* Cambridge: MIT Press.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms.* http://www.usf.edu/FreeAssociation/.

Pylkkänen, L.& McElree, B. (2006). The syntax-semantics interface: On-line composition of sentence meaning. In M. Traxler & M.A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed.). NY: Elsevier.

Rayner, K. (1998). Eye movements in reading and information processing: 20 yeas of research. *Psychological Bulletin, 124*, 372-422.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs: Prentice-Hall.

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility of eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1290-1301.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*, 445-476.

Shoben, E. (1991). Predicating and nonpredicating combinations. In P. J. Schwanenflugal (Ed.), *The psychology of word meanings* (pp. 117-135). Hillsdale, NJ: Erlbaum.

Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 327-342). Oxford, UK: Oxford University Press.

Staub, A., Rayner, K., Pollatsek, A., Hyönä, J., and Majewski, H. (2007). The time course of plausibility effects on eye movements in reading: Evidence from noun-noun compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*, 1162-1169.

Warren, T., & McConnell, K. (2007). Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. *Psychonomic Bulletin & Review*, *14*, 770-775.

Warren, T., McConnell, K., & Rayner, K. (under review). Effects of context on eye movements when reading about possible and impossible events.

WordNet: A Lexical Database for the English Language, http://wordnet.princeton.edu/.