NIH Public Access
Author Manuscript
J Exp Psychol Learn Mem Cogn. Author manuscript; available in PMC 2014 March 01.

Published in final edited form as:
J Exp Psychol Learn Mem Cogn. 2013 March ; 39(2): 487–501. doi:10.1037/a0028975.

Situational Context Affects Definiteness Preferences: Accommodation of Presuppositions

Charles Clifton Jr.
University of Massachusetts Amherst

Abstract

Four experiments used self-paced reading and eyetracking to demonstrate that readers are, under some conditions, sensitive to the presuppositions of definite vs. indefinite DPs (determiner phrases). Reading was faster when the context stereotypically provided a single possible referent for a definite DP or multiple possible referents for an indefinite DP than when context and DP definiteness were mismatched. This finding goes beyond previous evidence that definite DPs are processed more rapidly than indefinite DPs when there is a unique or familiar referent in the context, showing that readers are sensitive to the semantics and pragmatics of (in)definiteness. However, the finding was obtained only when readers had to perform a simple arithmetic task between reading a sentence and seeing a question about it. The intervening task may have encouraged them to process the sentence more deeply in order to form a representation that would persist while doing the arithmetic. The methodological implications of this observation are discussed.

Keywords

Indefinite noun phrase; definite noun phrase; accommodation; presupposition; context; eyetracking; self-paced reading; reading strategy

One widely-accepted psycholinguistic conclusion is that a text containing a definite noun phrase (DP)[i] is read and comprehended more rapidly than one containing an indefinite DP (Murphy, 1984). The effect is generally claimed to reflect the extra cognitive cost of introducing a new referent into discourse, as opposed to referring to an existing referent. The present research qualifies this claim without denying it. As discussed below, factors other than new vs. old may affect the ease of processing sentences with definite vs. indefinite DPs. The research reported here also advances a methodological claim: a reader's comprehension strategy can affect whether or not experimental techniques such as self-paced reading and eyetracking during reading can identify a subtle psycholinguistic effect, such as the effect that context has on the preference for a definite vs. indefinite DP.

Definite DPs (e.g., *the stove*) are commonly treated in the psycholinguistic literature as being anaphoric references to entities that were mentioned in the preceding discourse. In Murphy (1984) the noun in the DP, whose reading was facilitated by the definite determiner, was repeated from the earlier occurrence (even when the DP was indefinite). In related studies, Irwin, Bock, and Stanovich (1982) measured lexical decision times to words that continued a discourse, and found that preceding a previously-mentioned noun with the

Authors' address: Charles Clifton Jr., Department of Psychology, University of Massachusetts, Amherst, MA 01003 USA, cec@psych.umass.edu, 413 545 2653.

[i]I follow current linguistic usage, referring to the phrase consisting of a determiner plus a noun phrase, as a 'determiner phrase' (DP). This usage is motivated by both syntactic and semantic arguments.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

definite article *the* speeded the decision while the definite article did not consistently speed decisions to non-repeated words. Not all psycholinguistic research on definiteness has relied on literal repetition. A DP with a superordinate term (*the bird*) is easily processed as an anaphoric reference to a previously-introduced subordinate term, especially if the latter is a typical member of the category (*the robin*) (Garrod & Sanford, 1977). Similarly, psycholinguists have also studied 'bridging inferences' (Clark, 1975), in which previously mentioned entities or events support the inference that a potential referent existed in the situation. Sometimes this inference seems to be almost cost-free (as in *Keith was driving to London. The steering wheel…*) (Garrod & Sanford, 1982, 1994); sometimes it is costly (*We checked the picnic supplies. The beer was warm.)* (Clark & Haviland, 1977; Haviland & Clark, 1974). Burkhardt, 2006 (see also Schumacher, 2009) used ERP measures to demonstrate a similar cost to inferring an antecedent for a definite DP. This research measured ERPs while subjects read sentences following a context that mentioned a referent (*a speaker*), a situation that supported the inference of such a referent (*a lecture*), or a context that did not support such a referent (*Hannah*). A late positivity was observed to both definite and indefinite DPs (*the/a speaker*) following all contexts except for the context that explicitly mentioned the referent. This apparently reflected the processes required to introduce a new referent into the discourse, even in the case of a definite DP where the context supported the inference of the referent (a condition in which the second sentence was judged to be a highly plausible continuation of the first).

In the linguistic literature, definite DPs are most often discussed in terms of their requirements of uniqueness or of familiarity (or salience) (or, in some traditions, following Bertrand Russell, 1905, in terms of their role as quantifiers; we do not discuss this latter approach here). A definite DP is sometimes (in the tradition of Frege, 1892) claimed to require a referent that is unique (relative to some domain). Thus, if two safes are introduced into a discourse, it is said to be inappropriate to use the phrase *the safe* without modification. Alternatively, a definite DP can be claimed to require a referent that is familiar or somehow salient (Heim, 1982, 1983a; Lewis, 1979). Mentioning the referent in the previous discourse suffices to make it familiar (Roberts, 2003), but other factors such as situational salience or the structure of the discourse can make the referent appropriately salient (as a bridging inference can do). Often these two requirements are met simultaneously (although see Schwarz, 2009, who suggests that there may be two different definite articles, one for each kind of requirement, and argues that they appear as different phonological forms in some languages, e.g., the Fering dialect of German, spoken in the North Frisian island of Föhr).

Although it has not been presented as such, a good deal of psycholinguistic research can be viewed as examining the results of violating the uniqueness requirement. This research (e.g., Altmann, 1998; Altmann & Steedman, 1988; Crain & Steedman, 1985) has studied reading of discourses that first introduced one or two entities with the same word (e.g., one or two stoves) and then used a definite DP (*the stove*) followed by a relative clause or prepositional phrase (often one that is temporarily ambiguous syntactically). The relative clause is read more slowly, possibly showing garden-pathing effects, when only a single entity had been introduced than when two such entities had been. This has often been claimed to show that a relative clause or other postnominal modifier presupposes the existence of two or more potential antecedents, which must be distinguished. An alternative account (advanced by Clifton & Ferreira, 1989) is that a definite DP that violates the uniqueness requirement disrupts comprehension, and requires additional information (perhaps in the form of a following modifier) to satisfy the requirement.

While a definite DP that does not have a unique or familiar referent in preceding discourse can result in comprehension difficulty, in actual language use it is common to observe felicitous use of a definite DP without a preceding referent. Poesio and Vieira (1998) found

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

that less than half the occurrence of definite DPs in a sample of newspaper articles had an antecedent in the preceding discourse. Spenader (2002; cited in Frazier, 2005) and Frauaud (1990) have presented similar results. In many situations, readers and listeners appear to 'accommodate' (Lewis, 1979) the DP's presuppositional requirement by modifying their representation of the discourse to include the required presupposition. For instance, if you read a description of a situation in which a person goes to an appliance store to purchase a stove, your default representation of the situation presumably contains multiple stoves. The phrase *the stove* presupposes a single unique (or familiar) stove. You can accommodate this presupposition by assuming that the situation contains some particular stove that *the stove* can refer to, e.g. perhaps the stove that the person decided to purchase or the one that the salesperson was pushing. Heim (1983b), discussing how people accommodate missing presuppositions in various contexts, even suggests they deal with the problem 'effortlessly'. One question addressed in the current research is, when does such accommodation occur, and does it in fact carry a processing cost?

Indefinite DPs are at least as complex as definite DPs. Indefinite DPs have many distinct usages, including generic *A stove uses electricity*, nonreferential *I want a drink*, and various uses that may or may not refer, e.g. *I bought a car*). In some traditions (again following Russell), the indefinite article is viewed as a quantifier, but other, non-quantificational, analyses exist (e.g. Heim, 1983a).

In psycholinguistic research, indefinite DPs have mostly been studied as phrases that may introduce a new referent into the discourse. Clark and Haviland (1977) contrasted the effects of referential and nonreferential indefinite DPs on the reading of a following sentence containing a definite DP (e.g., *the alligator*). The definite DP sentence was read more slowly when the indefinite DP *an alligator* had previously been introduced in a nonreferential context (*Ed wanted an alligator for his birthday*) than when it had been introduced in a referential context (*Ed got an alligator for his birthday*). In the former case, the indefinite DP does not actually introduce a referent into the discourse model; readers apparently had to accommodate the presuppositions of the definite DP by assuming that Ed did get an alligator.

Various experiments have identified a processing cost to adding a referent to the discourse. For instance, Murphy (1984) found that reading was disrupted when a sentence like *Once his bicycle broke down* was followed by *Fortunately, he had a bicycle ready in time for the race*, compared to when *the bicycle* appeared in the second sentence. Presumably, if the bicycle mentioned in the second sentence was the same as that in the first sentence, the definite DP would have been used; since it was not, a second bicycle must be introduced, and that requires processing time. Schumacher (2009), who measured ERPs during sentence reading, observed a late positivity to all instances of indefinite DPs, even when context had explicitly introduced the noun of the DP. This presumably reflected the cost of introducing the new discourse referent denoted by the indefinite DP, a conclusion very similar to that made by Muprphy (1984).

To sum up, most psycholinguistic research on DP definiteness effects has concentrated on cases where the DP does refers to a previously-introduced discourse referent – cases that actually represent only a minority of observed uses of definite DPs (Poesio & Vieira, 1998, among others). A small amount of research (not all of which is available in published form) has examined how readers 'accommodate' the uniqueness and/or familiarity presuppositions of definiteness when discourse does not contain an antecedent to a DP. Frazier (2005) presented some evidence, stemming from undergraduate thesis work done by Evans (2005), that readers do make the kinds of accommodation that seem to be required by definite DPs. In one study, subjects read a sentence which described a protagonist either 'driving up to' or

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

'walking past' a row of toll booths. This sentence was followed by a sentence like *The tolltaker was asleep and drooling* vs. *A tolltaker was asleep and drooling.* A stops-making-sense task was used to assess the difficulty of processing the second sentence. The 'driving up to' context presumably makes one tollbooth, and thus one tolltaker, salient; the 'walking past' context does not. The sentences with a definite DP were accepted as sensible more often when they followed the 'driving up to' context that made one tolltaker salient than when they followed the 'walking past' context. The opposite was observed for sentences with an indefinite DP. Similarly, in a rating study involving only definite DPs, sentences with a definite DP were rated as more natural when they followed a context that presupposed the existence of a single referent (like the 'drove up to' context) than when they followed a context that required accommodation of such a presupposition (like the 'walked past' context).

Even though Evans's (2005) work provided judgment-based evidence that readers are intuitively sensitive to the difficulty of accommodating the presuppositions of a definite or an indefinite DP, several unpublished self-paced reading experiments conducted in the University of Massachusetts laboratories using materials like those Evans used failed to demonstrate any online differences in accommodation cost. For instance, subjects read sentences with a definite DP no faster than ones with an indefinite DP following a context that suggested the existence of a single referent (e. g, a small private beach….*the/a lifeguard*), and similarly, no differences were found following a context suggesting multiple potential referents.

The research to be reported here seeks evidence for the processing cost of accommodating the presuppositions of both definite and indefinite DPs. It extends Murphy's (1984) work by introducing situations in which there is stereotypically just a single instance (or multiple instances) of the entity referred to, without prior mention of the entity. For instance, a person's kitchen generally has a single stove. The uniqueness presupposition of the definite DP *the stove* requires the inference of a single stove, an inference that is supported by stereotypical knowledge of kitchens. In Garrod and Sanford's (1982) terms, the single stove would be in 'implicit focus.' However, use of the referential indefinite *a stove* in such a context would be pragmatically inappropriate given the availability of the definite DP. The use of the indefinite DP must be justified by accommodating its presupposition of multiple stoves. Such accommodation may incur a processing cost. Conversely, the presuppositions of an indefinite referential DP are supported by a context that stereotypically contains multiple instances of the entity referred to (e.g., multiple stoves in an appliance store), a context that does not support the presuppositions of a definite DP without added inferential work.

The experiments to be presented here also have a methodological component, motivated in part by the repeated failure (mentioned above) to demonstrate on-line effects of definiteness accommodation. This component will be described in detail in the introduction to Experiment 2, but to anticipate briefly, readers were required to do a secondary arithmetic task between reading a sentence and seeing a question about it. This task was intended to interfere with superficial memory of the form of the sentence and require more complete comprehension if the question was to be answered accurately (alternative possible mechanisms will be considered in the General Discussion).

Experiment 1

All four experiments to be reported had the same basic manipulation. The initial phrase of a sentence introduced a context that stereotypically had either a single instance of a to-be-mentioned object, or multiple instances of the object. Examples include someone's kitchen

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

vs. an appliance store preceding the mention of a stove; a person's desk vs. a classroom preceding mention of a chair; and a unicycle vs. a bicycle preceding mention of a tire. The sentence then continued by mentioning an entity either as a definite or an indefinite DP, as the direct object of the matrix sentence or the object of a preposition in the matrix verb phrase. All sentences were most naturally interpreted as containing a referential use of the DP, even in the case of indefinite DPs.[ii]

Following the analysis given in the Introduction, a definite DP would be appropriate for the context that suggested a single instance of the object, and an indefinite DP would be appropriate when the context suggested multiple instances. These cases will be referred to as the cases where definiteness 'matched the context.' Since the object referred to in the DP had not actually been mentioned in the discourse, the reader would have to accommodate its existence (or the existence of multiple instances), but accommodation would presumably be relatively straightforward and easy. More extensive and possibly more costly accommodation would be required when definiteness mismatched the context than when definiteness matched the context. In the case of a definite DP in a multiple-item context, presumably the reader would have to invent some reason to make one possible referent salient (e.g., in the examples given, the stove that the person wanted to buy, the chair that the person sat down in, the tire that needed repair). In the case of an indefinite DP in a single-item context, the context would have to be re-imagined to contain multiple instances: the kitchen would have to have multiple stoves, etc. Experiment 1 was designed to see whether the cost of this presumably more extensive accommodation would appear in a self-paced reading task.

Methods

Materials—Sixteen sets of four sentences were constructed, as illustrated in Table 1 (all appear in Appendix 1). Analysis regions are indicated by ^ symbols. The first region, Region 1, was not analyzed. The critical DP (Region 3) in each sentence was the last phrase of the initial clause of the sentence (generally a direct object), and appeared either as an indefinite (e.g., *a stove*) or a definite (*the stove*) DP. It was followed by a second clause or adjunct phrase, which served as the final region of the sentence (Region 4). Each sentence began with a phrase (Region 1) that introduced a situation that is stereotypically associated with either a single entity of the type referred to by the critical DP (e.g., *in the kitchen;* critical DP *the stove*) or multiple entities of this type (*in the appliance store*). Sentences with these two types of introductory phrases will be termed 'singular' and 'multiple' respectively. A comprehension question was made up for each set of sentences. For instance, the question of the 'kitchen' sentence used an an example, asked whether the protagonist was checking out something that he could cook with, or could clean with. The questions generally required only a fairly superficial understanding of the sentence, and their answers did not depend on the definiteness of the critical DP or the numerosity of its referent. All questions appear in Appendix 1.

The resulting 16 sentences were combined with 5 practice sentences and a total of 90 sentences from other experiments (testing sentences with comparative constructions, focus-inducing *only* phrases, quantifiers like *mostly*, and weak definites).

Subjects and procedures—Thirty-six University of Massachusetts undergraduates (in all experiments, approximately ¾ of whom were female), participating for course extra

[ii]The definite description usages in the materials could be classified in Hawkins' (1978) terms as 'larger situation uses,' in which the writer appeals to the reader's knowledge of entities that generally exist in the situations being talked about – although as Poesio and Vieira, 1998, note, classification of DPs is far from highly reliable.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

credit, were tested in individual half-hour sessions. A session began with brief written instructions, indicating that the subject should read naturally but be sure to understand what was read because each sentence would be followed by a question. The subject was to press the space bar on a computer keyboard to see the next phrase in a sentence, at which time the previous phrase would disappear. After the final phrase, a question appeared, with two answers, identified with the letters F and J, and the subject was to press the letter corresponding to the correct answer. The five practice sentences and questions were then presented, followed by an individually randomized presentation of all 106 items in the experiment. The program Linger (Rohde, 2003a) was used to present the instructions and the text and to record all answers and reading times.

Results

Accuracy—After one question that was inappropriate for the multiple context items was eliminated (it asked about a birthday party, which was mentioned in the singular but not the multiple conditions), subjects were accurate on 93% of the singular/indefinite and singular/definite items, and 92% of the multiple/indefinite and multiple/definite items. None of the differences approached significance (all $z < 1.0$) in a logistic linear mixed model analysis (Jaeger, 2008)

Reading times—Individual trial reading times for all trials, whether or not the question was answered correctly (a practice followed in all the experiments reported here) were analyzed as linear mixed models, using the program lmer (Bates & Sakar, 2007) under the R programming language (R Development Core Team, 2007). In each analysis, the fixed effect factors were context (singular vs. multiple) and definiteness (indefinite vs. definite). Sum coding (Crawley, 2007; also termed ‘contrast’ or ‘ANOVA-style’ coding) was used, so that the effect of each factor was evaluated at the mean of the other factors. Region length was centered and treated as a continuous predictor variable, adding to (but not interacting with) the other fixed effect variables. It had consistently large and significant effects on all measures, as is commonly found in self-paced reading; its effects will not be presented in the Results section or discussed further. The random factors were subjects and items. Models with random intercepts were initially compared with models that had random slopes involving the interaction of the context and definiteness. If the more complex model was significantly better than the simple model by a likelihood ratio test (Baayen, 2008), it was simplified by progressively removing random slope terms from the random factors as long as the more complex model was not significantly better than the simplified model (using a conservative significance level of $p < .10$). In all cases, the model with random intercepts proved to be the best and simplest model. A Monte Carlo Markov Chain procedure (Baayen, 2008) was used to estimate significance levels; thus, p(MCMC) values will be reported.

Individual trials with reading times < 200 ms were deleted (on the assumption that their contents were not actually read in such a short time). This resulted in the loss of less than 1% of all trials. Further, trials were deleted from analysis if they deviated by 3.0 SD or more from the values predicted by the model (standardized residuals > 3.0 SD; see Baayen, 2008). This resulted in the loss of 2% or fewer trials in each analysis.

Mean reading times for each region (after the 200 ms and the 3.0 SD standardized trimming) appear in Table 2. In Region 2, the subject and verb (plus preposition) of the matrix sentence, neither main effect (context, definiteness) approached significance ($F < 1.0$), but the interaction of the two factors was significant ($b = -28.8$, $t = 2.00$, $SE = 14.4$, $pMCMC = .046$). Sentences with a multiple context and an indefinite DP were read notably slowly. Since the manipulation of definiteness did not appear in Region 2, but only in Region 3, this result has to be spurious. Region 3, which contained the definite or indefinite DP, yielded no

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

significant effects (maximum t = 1.48, SE = 8.5, pMCMC > .14) or interaction (t = 1.38, SE = 8.5, pMCMC > .16). The apparent difference between indefinite and definite DPs that can be seen in Table 2 may simply reflect the greater length of the definite determiner, and was nonsignificant (F < 1) when region length was included as a predictor. The apparently-fast time for the multiple indefinite condition in Region 3 mirrors the spuriously slow time for this condition in Region 2, but the trial-by-trial correlation of times in Regions 2 and 3 was a negligible r = .006. There is no clear reason for this apparently aberrant pattern of slow then fast results (which did not appear in the next experiment). Finally, and most importantly given the results of Experiment 2, no effect approached significance in the spillover region, Region 4 (all t < 1.0). There was no sign of any faster reading when definiteness was appropriate for the context (singular/definite and multiple/indefinite; see the top panel of Figure 1).

Discussion

Experiment 1 quite simply failed to show that readers were sensitive to the presumed appropriateness of a definite vs. indefinite determiner for the situation under discussion. This is surprising, given people's intuitions when they are asked which of the sentence forms used in the experiment are more natural. Very consistently, people respond that the matching conditions are clearly preferred to the mismatching conditions. It is possible that readers are able to accommodate the presuppositions of a definite DP very quickly and easily (e.g., imagining a kitchen with two stoves, or an appliance store that had a stove they were particularly interested in). But even more likely is the possibility that subjects in Experiment 1 were reading with little comprehension, simply scanning the sentences, identifying the words in them, parsing them, and assigning a superficial analysis that was adequate to answer a question that immediately followed the sentence. Experiment 2 was designed to test this possibility.

Experiment 2

The second experiment used the same procedure and materials as Experiment 1, except that a simple arithmetic task was presented between the reading of the sentence and the presentation of the question. This procedure was based on unpublished work by Rohde (2003b), who measured self-paced reading performance in a situation where the sentence was followed by the presentation of a looping series of nonsense syllables which the subject had to repeat while reading and answering two questions about the sentence. Rohde reasoned that the questions that are typically asked in reading experiments are very shallow and can be answered by consulting the 'phonological buffer,' without requiring the reader to form a full thematic understanding of the sentence. Rohde reported slower reading by subjects who had this syllable-recitation secondary task than by subjects who did not, and more crucially, he found that the secondary task enhanced the effects of his substantive manipulations (subject- vs. object-extracted relative clauses, in embedded or right-branching position).

It is possible that the failure to find effects of definiteness in Experiment 1 was because the readers did not have to fully understand the sentences, but could answer the questions based on a superficial 'echo' of what they had just read. We return to an examination of this possibility below. In any event, Experiment 2 used a modification of Rohde's (2003b) technique to explore the possibility that increasing memory demands would result in fuller processing of the sentences and the appearance of the expected effects of the difficulty of accommodating the (in)definiteness presuppositions. Instead of directly engaging the speech production system, the experiment presented an arithmetic task between when the subject finished reading the sentence and when the question was presented. The assumption was that doing the arithmetic computation would occupy any phonological loop rehearsal and prevent

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

readers from consulting their 'echo' of the sentence. In order to answer questions accurately, readers were expected to have to comprehend the sentence to a 'deeper' level (Craik & Lockhart, 1972, and many following demonstrations of the effects of processing demands on memory). This, in turn, might force them to process the implications of the definite vs. indefinite determiner and its relation to the situation described in the sentence, which could possibly result in the appearance of the expected context-matching effect.

Methods

Experiment 2 differed from Experiment 1 only in that each sentence was immediately followed by a simple arithmetic task, adding or subtracting a one-digit number to or from a two-digit number (e.g., 38-5). The subject typed his or her answer in, following which a comprehension question appeared. The sentences and questions were identical to those used in Experiment 1. A different 36 University of Massachusetts undergraduates were tested. Testing procedures were the same as in Experiment 1, except for the mention of the arithmetic task that intervened between reading the sentence and seeing the question, plus the following text: "The point of the arithmetic problem is to erase your short-term memory of the sentence and encourage you to understand it when you read it, so you can answer the question accurately."

Results

Accuracy—Subjects were quite accurate on the arithmetic problems that followed each sentence (91.4% accurate for the problems that followed the 16 experimental sentences described here). After eliminating the question that was eliminated in Experiment 1, subjects were accurate on 96% of the comprehension questions of singular/indefinite items, 96% of the singular/definite items, 91%% correct on the multiple indefinite items, and 92% correct on the multiple definite items (all differences were nonsignificant; z always < 1.5, $p > 0.13$). Accuracy was closely comparable to that observed in Experiment 1, indicating that readers compensated for the intervening arithmetic task.

Reading times—Individual trial reading times were analyzed as linear mixed models, exactly as described for Experiment 1 (and again, models with random intercepts but not random slopes proved to be the most justified). Mean reading times for each region (after the 200 ms trimming and the 3.0 SD standardized residuals elimination) appear in the lower panel of Table 1. In Region 2, the subject and verb (plus preposition) of the matrix sentence, neither main effect (context, definiteness) nor the interaction approached significance ($t < 1.0$) The presumably-spurious interaction of context and definiteness observed in Experiment 1 did not appear in Experiment 2. Region 3, which contained the definite or indefinite DP, yielded no significant effects or interaction (maximum $t < 1.0$). The apparent difference between indefinite and definite DPs that can be seen in Table 2 once again was nonsignificant ($F < 1$) when region length was included as a predictor, and the deviantly-fast times for the multiple/indefinite condition observed in Experiment 1 did not appear here. Finally, and most importantly, while the main effects of context and definiteness in the spillover region, Region 4, were nonsignificant (all $t < 1.0$), the interaction of the two factors was significant ($t = 2.18$, $b = 122$, $SE = 55.9$, $pMCMC = .03$). Reading in the spillover region was faster when definiteness was appropriate to the situation (singular situation, definite, and multiple situation, indefinite) than when it was not. The effect can be seen in the bottom panel of Figure 1.

Discussion

The predicted effect of appropriateness of the determiner appeared in Experiment 2, while it was absent in Experiment 1. This seems to mean that, when subjects are induced to read

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

sentences well enough to retain them in memory during an intervening arithmetic task, they do note the situational appropriateness of the determiner and (if inappropriate) either accommodate presuppositions as required (by e.g. reinterpreting the singular situation to include multiple potential objects of the type mentioned) or pass it over after some reading disruption. Experiment 2 showed that intuitively clear and theoretically predicted pragmatic preferences can affect online processing when an adequate technique is used. There are interesting questions about why the technique had its effect, but these must largely be left for future research. It is of some interest nonetheless to note that reading overall appeared to be slower in Experiment 2 than in Experiment 1, while question-answering accuracy remained high. Mean total sentence reading time was 869 ms in Experiment 1 and 933 ms in Experiment 2. However, while this difference was highly significant in an analysis of variance with items as the random effect ($F(1,15) = 9.32$, $p < .001$), it was nonsignificant in the analysis by subjects ($F(1,70) < 1.0$); the between-subjects error sum of squares was extremely high, reflecting large individual differences in reading rate. Despite the nonsignificance of the latter effect, one might conclude that subjects in Experiment 2 read more carefully.

Other techniques could possibly result in obtaining the predicted match/mismatch effects. For instance, simply asking questions that required a more detailed representation of the situation described by the sentence might have produced the expected effects of definiteness (similar to what Swets, Desmet, Clifton, & Ferreira, 2009, found to be the case in sentences with ambiguous relative clause attachments). The questions used in the present experiments did not specifically require readers to create situation representations that specified the numerosity of possible referents (see Appendices 1 and 2). Asking some questions that did have such a requirement (e.g., *How many stoves were there in the kitchen? One ___ More than one ___*) might have yielded different results. We return to this question in the General Discussion.

The predicted effect appeared in the region following the region where the manipulation took place. It could be that the kind of adjustment a reader makes to inappropriate definiteness is somewhat delayed, appearing only after the reader has realized the presupposition violation. Alternatively, it could simply reflect the fact that effects often appear in a self-paced reading task in the spillover region (Mitchell, 1984, 2004). Experiments 3 and 4 are similar to Experiments 1 and 2, except that, instead of using self-paced reading, they measure readers' eye movements while they read. This technique has often been shown to be capable of showing processing effects on the region where the manipulation leading to the effect occurred, for both low-level manipulations involving word recognition (Rayner, 1998) and high-level manipulations involving interpretation (e.g., Rayner, Warren, Juhasz, & Liversedge, 2004). The technique has often been successful in showing effects of lexical, syntactic, and semantic processing difficulty without having to resort to imposing additional tasks on the reader, as was done in Experiment 2 (Clifton, Staub, & Rayner, 2007).

Experiment 3

Methods

Materials—The 16 sets of four sentences used in Experiments 1 and 2 were used in Experiment 3. They were slightly modified to make them appropriate for reading while eye movements were monitored. Most critically, material was added to the end of most sentences to ensure that the last fixation on a sentence (which could be terminated by a button press, not an eye movement) did not occur on a critical region. All materials appear in Appendix 2, which also shows the two-choice question asked of each item. A sample set of materials appears in Table 3.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

These 16 sentences were combined with eight practice sentences plus 120 sentences from other experiments plus fillers. The other experiments investigated effects of the stereotyped duration of events described, noun compounds, comparative structures, and weak definite noun phrases. A two-choice question followed each of the 16 sentences used in Experiment 3, plus 56 of the remaining 120 sentences.

Subjects and Procedures—Data from 38 University of Massachusetts undergraduates, working for extra course credit, are presented. One additional subject's data had to be eliminated because of excessive track losses. The experimental sentences were presented as four counterbalanced lists, with four sentences appearing in each of the four versions illustrated in Table 3 in each list. The context (single vs. multiple) and definiteness (indefinite vs. definite) were crossed within-item and within-subject factors.

The 16 items that a subject saw appeared in individually randomized order, intermixed randomly with the remaining 120 items (and following the practice items). Subjects were instructed to read each sentence for comprehension, and then press a key on a game controller to see and answer the question that appeared after some items. They also pressed the left or the right key on the game controller to indicate whether the correct answer was on the left or the right of the screen.

Subjects' eye movements were recorded while reading the sentence. They were seated 55 cm from a CRT monitor that displayed the sentences and their eye movements were recorded using an EyeLink 1000 (SR Research, Toronto, Ontario, Canada) eyetracker controlled by a PC running the University of Massachusetts EyeTrack software (http://www.psych.umass.edu/eyelab). Eye position was sampled at 1 kHz. Although participants viewed the stimuli with both eyes, only data from the right eye were collected. The eyetracker was calibrated and then the experiment began. A trial was initiated when the participant fixated on a box located where the first letter of the question would appear. All items were presented on a single line, in a proportional-width font (Ariel 16 pt.). Calibration was checked on each trial by checking the alignment with a single centered target. Unacceptably large errors led to recalibration. The entire experiment took 45 minutes or less.

Results

Overall accuracy on the comprehension questions (with the inappropriate question mentioned in Experiment 1 replaced by an appropriate one) averaged 93% correct for singular/match items and singular mismatch items, 91% for multiple/match items, and 92% for multiple/mismatch items, with no significant differences across conditions, indicating good comprehension. The eye movement data were cleaned of track losses, blinks, and long fixations over 800 ms using the program EyeDoctor (http://www.psych.umass.edu/eyelab). Following Rayner, Sereno, Morris, Schmauder, & Clifton (1989), we assume that short fixations, i.e., those under 80 ms, do not contribute useful information to the reader. Thus, short fixations were incorporated into the nearest neighboring fixation within three characters; otherwise, they were deleted. Fixations greater than 800 ms were also eliminated, on the assumption that these did not represent normal acquisition of information from the text (Rayner et al., 1989). Finally, trials with blinks or track losses on the critical word were eliminated, 8.2% of all trials.

The sentences were divided into several regions for analysis purposes, as illustrated in Table 3. The main region of interest was Region 3, the definite or indefinite DP, but the following region was also analyzed to identify any delayed or spillover effects. The informative eyetracking measures are presented in Table 4: first-pass time (the summed fixation durations from first entering a region until first leaving the region); second pass time (the

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

summed duration of all fixations in a region); first pass regressions out (the proportion of trials on which a first-pass fixation in a region was followed by a regression to an earlier region of the sentence); and number of fixations (the mean number of fixations made in a region from first entering to first leaving it). Except for the second pass time measure, trials on which no fixation was made in a region were eliminated from analyses of that region, on the assumption that a zero fixation time would not really mean that the region was processed without looking at it. See Rayner et al., 1989, for descriptions of these measures.

Each measure except for second pass time was analyzed as a linear mixed model as in Experiment 1, with sum coding of the fixed effect factors (context and definiteness) and with region length as a centered additive predictor variable. As in the self-paced reading experiments, centered length had consistently large and significant effects on all measures, as is essentially always found in eyetracking (Rayner, 1998); its effects will not be presented in the Results section. The random factors were subjects and items. Models with different random factor structures were evaluated and selected as described in connection with Experiment 1. Again, random intercept models were the preferred models, never being significantly worse ($p < .10$) than more complex models. Proportions of regressions out were analyzed using logistic regression (Jaegger, 2008). Because second pass time data severely violate normality at the individual trial level (the modal second pass duration is often zero ms), by-subject and by-item average second pass times were analyzed as F1 and F2 ANOVAS, using the ez-Anova function available in the R package ‘ez.’ Region length was not entered as a factor in these analyses; computing second pass time as a linear function of region length on a trial-by-trial basis (as would be needed to analyze residualized times as required in at least the by-subjects analyses) would not be legitimate, given the preponderance of zero-duration times.

As in Experiments 1 and 2, Region 1 was not analyzed, nor was the last region of the sentence, Region 5. Analyses of first pass time will be presented first, as they proved to be most informative in Experiment 4, to be reported next. Neither the effect of context nor the effect of definiteness, nor their interaction approached significance in Region 2 (all $t < 1.0$). In Region 3, there were nonsignificant hints of effects of context and definiteness ($t = 1.62$, $b = 11.3$, $SE = 7.0$, $pMCMC < .11$, and $t = 1.71$, $b = -12.7$, $SE = 7.4$, $p < .09$, respectively; DPs following multiple contexts tended to be read more slowly than DPs following singular contexts, and definite DPs tended to be read more slowly than indefinite DPs, despite the inclusion of region length as a predictor. There was, however, no sign of an interaction ($t < 1.0$). Mean first pass reading times, unadjusted for the greater length of definite than indefinite DPs, appear in the top panel of Figure 2. In Region 4, first pass times were longer following singular than following multiple contexts ($t = 2.05$, $b = -17.9$, $SE = 8.7$, $p < .05$), but neither the effect of definiteness nor the interaction of context and definiteness was significant ($t = 1.63$, $b = -14.3$, $SE = 8.7$, $pMCMC > .10$, and $t = 1.50$, $b = -13.1$, $SE = 8.7$, $pMCMC > .14$, respectively. The hint of an interaction was not what was predicted on the assumption that reading would be disrupted when definiteness was inappropriate for the context. Rather, reading seemed to be marginally faster than expected in the multiple context, definite condition.

Few significant differences were observed in the other measures. Considering the proportion of trials on which a first-pass regression out occurred, there was only one significant effect in one region: In region 3, there were more regressions out of indefinite than definite DPs ($z = 2.49$, $b = -0.36$, $SE = 0.14$, $p < .02$). However, the interaction of context and definiteness was nonsignificant ($z = 1.01$). No other effect in any region reached significance beyond the $p = .16$ level. Number of fixations effects did not approach significance in any region (largest $t = 1.28$, $b = 0.036$, $SE = 0.028$, $p > .20$). The only hint of a significant effect in second-pass time appeared in the Region 2 interaction of context and definiteness ($F1(1,37)$

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

= 3.40, η^2 = 0.023, p < .08, but F2(1,15) = 1.70, η^2 = 0.030, p > .20). Besides being nonsignificant, this numerical effect goes in the direction opposite to that expected: second pass times tended to be longer in the conditions where definiteness was appropriate for the context, rather than being faster.

Discussion

Experiment 3 provided eyetracking data that were consistent with the self-paced reading results of Experiment 1. There was no suggestion that sentences with a definite DP were read faster in a context that supported a unique, identifiable referent than in contexts that suggested multiple referents, nor were sentences with indefinite DPs read faster in the latter than the former context. However, as suggested by Experiment 2, this result should perhaps not be taken as disconfirmation of the hypothesis that comprehension would be facilitated when definiteness is appropriate for the context. Experiment 2 indicated that separating the self-paced reading of a sentence from a question about the sentence by requiring a distracting arithmetic task did result in observing the predicted effect, presumably by encouraging more careful reading and deeper comprehension of the sentence. Experiment 4 was designed to determine whether a similar effect could be observed using eyetracking, and if so, to determine how quickly it would appear in reading. As indicated earlier (Mitchell, 2004), self-paced reading results often appear after the region that contains the manipulation, in a ‘spillover’ region. Eyetracking measures often indicate the faster appearance of an effect, giving a more accurate picture of just where in reading some processing occurred.

Experiment 4

Methods

The sixteen sets of sentences used in Experiment 3 were combined with the same practice items plus 80 sentences from other experiments (including experiments on restrictive vs. nonrestrictive relative clauses, interpretation of reflexive anaphors, and comparison of noun and adjective descriptors) and filler sentences. Each sentence was followed by a simple arithmetic problem, presenting one 2-digit number, a plus or minus sign, and a 1-digit number, followed on the next line by two possible results, one correct, one in error, half to the left of the screen and half to the right. Subjects were to pull the trigger on the game controller under the answer they thought was correct.

A total of 52 University of Massachusetts undergraduates were tested for course extra credit. Seven could not be tracked, and four additional subjects were eliminated for having too many track losses, yielding a total of 41 subjects. Apart from the arithmetic task that intervened between reading the sentence and seeing and answering the question, the only differences from Experiment 3 were that the items were presented in a fixed-width Monaco font, and the instructions that were read to the subjects described the arithmetic task and included the sentence “You’ll soon find that you have to read the sentence quite carefully to be able to remember what it said when the question appears.”

Results

Accuracy in answering the arithmetic questions averaged 91%. The questions about the sentences were answered correctly 94% of the time in the singular/match, singular mismatch, and multiple/match conditions, and 92% in the multiple mismatch condition. The values did not differ significantly and were closely comparable to Experiment 3.

The eyetracking data, which appear in Table 5, were analyzed just as in Experiment 4 (and again, the models selected for analyzing the reading time data all had random intercepts of

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

subjects and items, except of course for the analyses of second pass time). Data from first pass time will be presented first, because they provided clear evidence about the hypothesis being tested. In Region 2, before the definiteness manipulation took place, there were nonsignificant hints of effects of definiteness and its interaction with context (t = 1.64, b = 17.2, SE = 10.5, pMCMC = .11, and t = 1.49, b = −15.6, SE = 10.5, pMCMC = .14, respectively). The trend in the data suggested slow reading times in the singular, definite DP condition, quite contrary to any expectations. However, in Region 3, the DP region, the expected interaction did appear (t = 2.11, b = 13.9, SE = 6.6,, pMCMC = .04). It can be seen in the bottom panel of Figure 2: reading was speeded when the definiteness of the DP was appropriate for the context. No effects approached significance in the spillover region (where the interaction had been seen in the self-paced reading Experiment 2) (all $p > .20$).

The other measures added little to the evidence. The only hint of an effect in the proportion of regressions out was the slightly elevated frequency of regressions out of Region 2 in the singular context condition (z = 1.55, b = −0.25, SE = 16.4, p = 0.12). No other region produced an effect (all $p > .18$). Number of fixations provided suggestive but only marginally significant evidence for speeded reading in Region 3 when the determiner matched the context. The same pattern seen in first pass time appeared to be present in number of fixations, but was not fully significant (t = 1.81, b = 0.052, SE = 0.029, pMCMC = .07). No other effects were significant in the analyses of number of fixations. Second pass times had a nonsignificant suggestion of being longer in Region 2 in the singular context conditions and the indefinite DP conditions ($F1(1,40) = 4.20$, $\eta^2 = 0.015$, $p < .05$ and $F2(1,15) = 3.81$, $\eta^2 = 0.059$, $p < .07$, and $F1(1,40) = 3.59$, $\eta^2 = 0.011$, $p < .07$ and $F2(1,15) = 2.61$, $\eta^2 = 0.029$, $p < .13$, respectively). The interaction F was < 1.0. Second pass time in Region 3 also seemed to be greater for indefinite than definite DPs ($F(1,40) = 4.95$, $\eta^2 = 0.015$, $p < .04$, but $F2(1,15) = 2.69$, $\eta^2 = 0.017$, $p < .13$), a tendency that may have been greater for multiple than for singular contexts, but the interaction approached significance only by items ($F1(1,40) = 0.65$, $\eta^2 = 0.003$, $p > .40$; $F2(1,15) = 3.74$, $\eta^2 = 0.014$, $p = .07$). Second pass time in Region 4 actually appeared to be relatively long in the conditions where DP definiteness matched the context, but the interaction was not significant ($F1(1,40) = 2.96$, $\eta^2 = 0.011$, $p > .09$; $F2(1,15) = 2.26$, $\eta^2 = 0.026$, $p > .15$).

Discussion

While the joint effect of context and definiteness was not overwhelmingly strong, it did appear in the initial reading of the critical DP. The interaction was significant for first pass reading time, and marginal for number of fixations. However, any difficulty was resolved during the reading of the critical region, and was not manifested in the probability of going back to reread previous material or the time spent re-reading material. Taking Experiment 2 into account as well, it does appear that readers are sensitive to the contextual appropriateness of the determiner used in a DP.

However, this sensitivity is not a given, at least if readers can content themselves with a shallow understanding of the text they read. It appears that they have to be encouraged to process the text deeply enough to survive a brief distraction before having to answer a question about the text. One point in support of this interpretation comes from comparing the second pass times in Experiments 3 and 4. Although the experiments were not designed to be directly compared (and such comparison is not the point of this research), it is interesting to note that subjects seemed to spend much more time re-reading in Experiment 4 than in Experiment 3. The mean second pass times, pooled over regions and conditions, were 97 ms in Experiment 3, and 150 ms in Experiment 4. While the comparable difference between Experiments 1 and 2 was significant only by items, the difference between Experiments 3 and 4 was fully significant ($t1(77) = 2.22$, $p < .03$; $t2(15) = 6.46$, $p < .001$).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

We suspect subjects in Experiment 4, in addition to processing the sentences more deeply initially, spent more time re-reading and rehearsing the sentences to ensure memory.

General Discussion

The reading time data from Experiments 2 and 4 support linguistic analyses of the presuppositions of definite and indefinite DPs. When the situational context does not stereotypically support these presuppositions, reading can be slowed. The fact that reading was slowed on the DP itself in the eyetracking Experiment 4 indicates that the presuppositional requirements of a DP can be identified quickly, comparable to how quickly a syntactic or semantic garden-path can be identified (Clifton et al., 2007). This finding indicates that it is too simple to say (e.g.) that definite DPs are necessarily easier to process than indefinite DPs (a tempting conclusion from Murphy, 1984, and Irwin et al., 1982). Presuppositionally appropriate DPs are easier to process than presuppositionally inappropriate ones.

The current data cannot unambiguously determine whether the slowed reading reflected time taken to accommodate the presuppositions, or simply disruption triggered by noting that presuppositions had not been met. There are reasons to favor the former alternative. Presumably some accommodation of presuppositions was required even in the conditions where discourse context matched definiteness. The referent of the definite DP had not been previously mentioned, nor had the alternative possible referents of an indefinite DP. As Schumacher's (2009) ERP data indicated, a reader must introduce the referent into the discourse when it had not previously been introduced, even when the discourse clearly supported the inference of the required referent. It seems likely that prior mention of the referent of a definite DP (e.g., a stove) in the present materials would have resulted in faster reading than was observed when context matched definiteness (cf. Murphy, 1984). Assuming that reading was slowed in the matching conditions relative to such a prior-mention baseline, reading was even more slowed in the conditions where the presuppositions of the DP were not put into 'implicit focus' (Garrod & Sanford, 1982) by the context. It seems very reasonable to suppose that this differential slowing reflected the differential difficulty of accommodating the DP's presuppositions, and thus that readers did typically make the required accommodation.

An additional reason to suspect that the slowed reading meant that readers did accommodate the presuppositions comes from the observation that reading was slowed in the experiments that introduced the arithmetic task between reading a sentence and seeing a question about it, compared to the experiments without the arithmetic task. It is very tempting to assume that the requirement to maintain a sentence in memory while doing the arithmetic induced readers to comprehend the sentence more deeply. This assumption is based on copious evidence that more extensive or deeper processing leads to better memory, and the assumption that while a very superficial reading of the sentence would leave a trace adequate to answer an immediately-presented question, such a trace would not consistently survive the arithmetic computation. Such deeper comprehension of a sentence would very reasonably include a representation of the situation described that met the presuppositions of the DP, and when the situation had non-stereotypical properties (e.g., a kitchen with multiple stoves, a store with one singled-out stove), it would be more difficult to create.

To be sure, the present experiments leave many questions unanswered. Just why the arithmetic task had its effect is uncertain. Did the effort required by the computation involve subvocal activity, interfering with a phonological representation of the sentence and encouraging fuller comprehension? There is no direct evidence in the present research that phonological representations were involved (although the results reported by Rohde, 2003b,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

whose secondary task was designed explicitly to interfere with such a representation, suggests that they were). Or did the anticipated need to do the arithmetic simply occupy some processing resources, which would have otherwise been adequate to comprehend the sentence and even do accommodation of presuppositions during the time required simply to read the words? Would very different manipulations have produced the same effect? For instance, as discussed earlier, asking questions that forced attention to the parts of a sentence that are directly relevant to the DP's presuppositions might have increased sensitivity to them (cf. Swets et al., 2009). Alternatively, putting greater emphasis on the setting (e.g., the kitchen in the example sentence used earlier) by describing it in a separate sentence rather than a sentence-initial prepositional phrase might have increased the accommodation effects.

Another unanswered question is, why did the present manipulation require the secondary task to produce an effect on reading time? Eyetracking and self-paced reading have effectively shown many lexical, syntactic, and even semantic and pragmatic effects without the need for a secondary task (see Mitchell, 2004, and Clifton et al., 2007, for reviews). For instance, semantic anomaly can have immediate effects on eye movement measures (Rayner, Warren, Juhasz, & Liversedge, 2004) without any need for an intervening task. In addition, while semantic coercion (e.g. treating reference to *a book* as if it were a reference to an event rather than an entity, as in *I started the book*) may not have immediate effects, its effects appear quickly (Traxler, Pickering, & McElree, 2002). Similarly, there are reports that pragmatic processing difficulty appears in the on-line reading record. Pronominal reference to a non-topical and not easily accessible referent clearly slows reading (Ehrlich & Rayner, 1983). The presuppositional requirements of focus particles like *only* and *even* seem to affect reading speed (Filik, Paterson, & Liversedge, 2009).

Perhaps the difficulty of showing on-line effects of the presuppositions of a DP can be traced to the fact that a relatively complete situational model must be built to accommodate them. A new discourse referent can be introduced into a simple situational model by either an indefinite or an unheralded definite DP. Only if the implications of the model are examined – does a kitchen have multiple stoves? – will the need to accommodate the DP's presuppositions be apparent. It is plausible that, without the demand to construct a stable and persisting representation of the sentence in Experiments 1 and 3, readers would not bother to build an appropriately complete situational model. This is, perhaps, particularly likely given that many of the experimental sentences began with novel definites (e.g., *The student walked behind the desk,*). The fact that these definite DPs were unsupported could discourage readers from creating an explicit situational model. This line of reasoning suggests that it might be possible to demonstrate accommodation effects in discourses that provided better support for a rich situational model.[iii]

Returning to the present data, the claim is that readers may have understood a sentence well enough to answer a question about it without building a relatively complete discourse model, except when their memory for the sentence was challenged by the need to complete the intervening arithmetic task. In this latter situation, they may have attempted to integrate the content of the sentence into a coherent model of the discourse, in which case they would evaluate and attempt to accommodate the presuppositions of definite and indefinite DPs, processes that appear in the on-line reading record.

Acknowledgments

I would like to thank Lyn Frazier and Adrian Staub for reading and commenting on earlier versions of this paper, and Seda Kan, Meg Grant, Angela Pazzaglia, Erin Ackley, Adina Galili, Brittany McArdle, Brittany Stepton,

[iii]I thank Lyn Frazier for this suggestion

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Morgan Mendes, Maria Bonilla, and Zachary Waegell for assistance in collecting the data reported here. This project was supported in part by Grant Number HD18708 from NICHD to the University of Massachusetts. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of NICHD or NIH.

References

Altmann GTM. Ambiguity in sentence processing. Trends in Cognitive Sciences. 1998; 2:146–152. [PubMed: 21227111]

Altmann G, Steedman M. Interaction with context during human sentence processing. Cognition. 1988; 30:191–238. [PubMed: 3215002]

Baayen, RH. Analyzing linguistic data. Cambridge: Cambridge University Press; 2008.

Bates, DM.; Sakar, D. lme4: Linear mixed effects models using S4 classes (Version R package version 0.998875–6). 2007.

Burkhardt P. Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. Brain and Language. 2006; 98:159–168. [PubMed: 16725188]

Clark, HH. Bridging. In: Schank, R.; Nash-Webber, BL., editors. Theoretical Issues in Natural Language Processing. Association for Computational Linguistics; 1975. p. 169-174.

Clark, HH.; Haviland, SE. Comprehension and the given-new contract. In: Freedle, R., editor. Discourse comprehension and production. Norwood, N.J: Ablex Publishers; 1977. p. 1-40.

Clifton C Jr, Ferreira F. Ambiguity in context. Language and Cognitive Processes. 1989; 4:SI 77–104.

Clifton, C., Jr; Staub, A.; Rayner, K. Eye movements in reading words and sentences. In: Gompel, RV.; Fisher, M.; Murray, W.; Hill, RL., editors. Eye movement research: Insights into mind and brain. New York: Elsevier; 2007. p. 341-371.

Craik FIMR, Lockhart S. Levels of processing: A framework for memory research. Journal of Verbal Learning and Verbal Behavior. 1972; 11:671–684.

Crain, S.; Steedman, M. On not being led up the garden path: The use of context by the psychological parser. In: Dowty, D.; Kartunnen, L.; Zwicky, A., editors. Natural language parsing. Cambridge: Cambridge University Press; 1985. p. 320-358.

Crawley, MJ. The R Book. Chichester, England: John Wiley & Sons; 2007.

Ehrlich K, Rayner K. Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. Journal of Verbal Learning and Verbal Behavior. 1983; 22:75–87.

Evans, W. Unpublished undergraduate thesis. University of Massachusetts; Amherst, MA: 2005. Small worlds of discourse and the spectrum of accommodation.

Filik R, Paterson KB, Liversedge SP. The influence of only and even on online semantic interpretation. Psychonomic Bulletin & Review. 2009; 16:678–683. [PubMed: 19648452]

Fraurud K. Definiteness and the processing of noun phrases in natural discourse. Journal of Semantics. 1990; 7:395–433.

Frazier, L. The big fish in a small pond: Accommodation and the processing of novel definites. 2005. Retrieved December 21, 2012, from OSU Pragmatics Initiative, http://ling.osu.edu/pragwiki/index.php/The_Big_Fish_in_a_Small_Pond

Frege G. Über Sinn und Bedeutung. Zeitschrift für Philosophie und philosophische Kritik. NF 100:25–50.

Garrod S, Sanford A. Interpreting anaphoric relations: The integration of semantic information while reading. Journal of Verbal Learning and Verbal Behavior. 1977; 16:77–90.

Garrod S, Sanford A. The mental representation of discourse in a focused memory system: Implications for the interpretation of anaphoric noun phrases. Journal of Semantics. 1982; 1:21–41.

Garrod, SC.; Sanford, AJ. Resolving sentences in a discourse context: How discourse representation affects language understanding. In: Gernsbacher, MA., editor. Handbook of Psycholinguistics. San Diego: Academic Press; 1994.

Haviland SE, Clark HH. What's new? Acquiring new information as a process in comprehension. Journal of Verbal Learning and Verbal Behavior. 1974; 13:512–521.

Hawkins, JA. Definiteness and indefiniteness. London: Croom Helm; 1978.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Heim, I. Unpublished Doctoral dissertation. University of Massachusetts; 1982. The semantics of definite and indefinite noun phrases.

Heim, I. File change semantics and the familiarity theory of definiteness. In: Bauerle, R.; Schwarze, C.; Stechow, Av, editors. Meaning, Use, and the Interpretation of Language. Berlin: de Gruyter; 1983a.

Heim, I. On the projection problem for presuppositions. In: Portner, P.; Partee, BH., editors. Formal Semantics: the Essential Readings. Oxford: Blackwell; 1983b. p. 397-405.

Irwin DE, Bock JK, Stanovich K. Effects of information structure cues on visual word processing. Journal of verbal Learning and verbal Behavior. 1982; 21:307–325.

Jaeger TF. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. Journal of Memory and Language. 2008; 59:434–446. [PubMed: 19884961]

Lewis D. Scorekeeping in a language game. Journal of Philosophical Logic. 1979; 8:339–359.

Mitchell, DC. An evaluation of subject-paced reading tasks and other methods_of investigating immediate processes in reading. In: Kieras, DE.; Just, MA., editors. New Methods in Reading Comprehension Research. Hillsdale, N.J: Erlbaum; 1984.

Mitchell, DC. On-line methods in language processing: Introduction and historical review. In: Carreiras, M.; Clifton, C., Jr, editors. The on-line study of sentence comprehension: Eyetracking, ERPs, and beyond. Brighton, UK: Psychology Press; 2004.

Murphy GL. Establishing and accessing referents in discourse. Memory & Cognition. 1984; 12:489–497. [PubMed: 6521651]

Poesio M, Vieira R. A corpus-based investigation of definite description use. Computational Linguistics. 1998; 24:183–216.

Rayner K. Eye movements in reading and information processing: 20 years of research. Psychological Bulletin. 1998; 124:372–422. [PubMed: 9849112]

Rayner K, Sereno S, Morris R, Schmauder R, Clifton C Jr. Eye movements and on-line language comprehension processes. Language and Cognitive Processes. 1989; 4:SI 21–50.

Rayner K, Warren T, Juhasz BJ, Liversedge SP. The effect of plausibility on eye movements in reading. Journal of Experimental Psychology: Learning, Memory and Cognition. 2004; 30:1290–1301.

Roberts C. Uniqueness in definite noun phrases. Linguistics and Philosophy. 2003; 26:287–350.

Rohde, D. Linger: A flexible program for language processing experiments. 2003a. Retrieved from http://tedlab.mit.edu/~dr/Linger/

Rohde, D. Assessing first-pass comprehension of relative clause sentences. Poster presented at 16th Annual CUNY Conference on Human Sentence Processing; 2003b.

Russell B. On denoting. Mind. 1905; 59:320–344.

Schumacher, PB. Definiteness marking shows late effects during discourse processing: Evidence from ERPs. In: Lalitha Devi, S.; Branco, A.; Mitkov, R., editors. Anaphora Processing and Applications, Lecture Notes in Artificial Intelligence. Vol. 5847. Heidelberg: Springer; 2009. p. 91-106.

Schwarz, F. Unpublished PhD Dissertation. Department of Linguistics, University of Massachusetts; Amherst, MA: 2009. Two types of definites in natural language.

Spenader, J. Unpublished PhD Dissertation. Computational Linguistics, Stockholm University; Stockholm, Sweden: 2002. Presupposition in spoken discourse.

Swets B, Desmet T, Clifton C Jr, Ferreira F. Underspecification of syntactic ambiguities: evidence from self-paced reading. Memory & Cognition. 2008; 36(1):201–216. [PubMed: 18323075]

Traxler M, Pickering MJ, McElree B. Coercion in sentence processing: Evidence from eye movements and self-paced reading. Journal of Memory and Language. 2002; 47:530–548.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Appendix 1. Materials used in Experiments 1 and 2, single context then multiple context

Indefinite/definite articles separated by /. Underscores connect words presented as a single region; @ indicate analysis regions. Questions with alternative answers, separated by space, follow the pairs of items, marked by !.

1. In_the_kitchen, Jason_checked_out@1 a/the_stove@2 very_carefully.

   In_the_appliance_store, Jason_checked_out@1 a/the_stove@2 very_carefully.

   Jason was checking out something {that_he_could_cook_with that_he_could_clean_with}

2. The_student_walked_behind_the_desk and_sat_down@1 in_a/the_chair@2 and_sighed.

   The_student_walked_into_the_classroom and_sat_down@1 in_a/the_chair@2 and_sighed.

   ! The student was {studying sitting_down}

3. In_her_bathroom_at_home, Sally_cleaned@1 a/the_sink@2 with_a_bristle_brush.

   In_the_Campus_Center_women's_room, Sally_cleaned@1 a/the_sink@2 with_a_bristle_brush.

   ! What was Sally doing with the sink? {cleaning_it using_it}

4. At_his_granddaughter's_birthday_party, Chuck_admired@1 a/the_cake@2 and_let_people_know_it.

   In_the_fancy_bakery, Chuck_admired@1 a/the_cake@2 and_let_people_know_it.

   ! Whose birthday party was it? {Chuck's_daughter's Chuck's_granddaughter's}

5. In_her_living_room, Susan_tried_out@1 a/the_hi-def_television@2 but_was_not_pleased.

   In_the_electronics_store, Susan_tried_out@1 a/the_hi-def_television@2 but_was_not_pleased.

   ! Susan was checking out {a_radio a_TV}

6. As_part_of preparing_his_unicycle, Billy_checked@1 a/the_tire@2 for_a_puncture.

   As_part_of preparing_his_bicycle, Billy_checked@1 a/the_tire@2 for_a_puncture.

   ! Billy was fixing his {cycle car}

7. In_her_basement, Becky_heard_a_bad_noise coming_from@1 a/ the_washing_machine@2 and_started_worrying.

   In_the_laundromat, Becky_heard_a_bad_noise coming_from@1 a/ the_washing_machine@2 and_started_worrying.

   ! What probably needed fixing? {a_dryer a_washer}

8. Sitting_in_front_of_his_PC, Sam_typed_his_name@1 on_a/the_keyboard@2 and_listened_to_the_clicks.

   Standing_in_the_aisle_at_Staples, Sam_typed_his_name@1 on_a/the_keyboard@2 and_listened_to_the_clicks.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

! What was Sam checking out? {a_computer_part a_TV_part}

9. Seeing_his_locker, Sandy_tried_to_open@1 a/the_door@2 but_failed.

Seeing_a_row_of_lockers, Sandy_tried_to_open@1 a/the_door@2 but_failed.

! What did Sandy try to open? {a_door a_window}

10. Picking_up_her_hamburger, Patty_noticed_mold@1 on_a/the_bun@2 and_was_disgusted.

Working_the_line_at_McDonalds, Patty_noticed_mold@1 on_a_bun@2 and_was_disgusted.

! What did Patty notice on the bun? {mold a_bug}

11. Playing_near_the_pine_tree, little_Billy_hid_behind@1 a/the_trunk@2 and_was_very_quiet.

In_the_grove_of_pine_trees, little_Billy_hid_behind@1 a/the_trunk@2 and_was_very_quiet.

! What did Billy hide behind? {a_trunk a_car}

12. Standing_in_front_of_the_Congregational_church,

Annette_noticed_some_missing_shingles@1 on_a/the_steeple@2 and_was_sad.

Standing_in_front_of_the_row_of_churches, Annette_noticed_some_missing_shingles@1 on_a/the_steeple@2 and_was_sad.

! What did Annette notice was missing? {some_shingles some_plants}

13. Looking_at_the_parked_car, Fred_thought_he_recognized@1 a/the_driver@2 but_realized_he_was_wrong.

Looking_at_the_line_of_cabs, Fred_thought_he_recognized@1 a/the_driver@2 but_realized_he_was_wrong.

! What did Fred think he recognized? {a_driver a_friend}

14. Putting_the_single_flower into_the_vase, Abby_had_to_trim@1 a/the_stem@2 because_the_vase_was_short.

Putting_the_bunch_of_flowers into_the_vase, Abby_had_to_trim@1 a/the_stem@2 because_the_vase_was_short.

! What did Abby put into the vase? {a_flower a_plant}

15. Standing_on_the_porch_of_his_house, Joe_put_an/the_oddshaped_package@1 into_a_mailbox@2 and_walked_away.

Standing_in_the_lobby_of_the_apartment_complex, Joe_put_an/the_oddshaped_package@1 into_a_mailbox@2 and_walked_away.

! What was Joe standing on? {the_porch the_sidewalk}

16. Standing_at_his_kitchen_counter, Tom_reached_for@1 a/the_coffeepot@2 and_poured_a_cup.

Standing_at_the_Dunkin_Donuts_self-serve_counter, Tom_reached_for@1 a/the_coffeepot@2 and_poured_a_cup.

! What did Tom reach for? {coffeepot newspaper}

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Appendix 2: Materials used in Experiments 3 and 4; single context then multiple context

Indefinite/definite articles separated by /. ^ indicate analysis regions. Questions with alternative answers follow the pairs of items.

1. In the kitchen,^ Jason checked out^ a/the stove^ very carefully^ before using it.^

   In the appliance store,^ Jason checked out^ a/the stove^ very carefully^ before using it.^

   Jason was checking out something…

   that he could cook with. that he could clean with.

2. The student walked behind the desk and^ sat down in^ a/the chair^ and sighed^ out loud.^

   The student walked into the classroom and^ sat down in^ a/the chair^ and sighed^ out loud.^

   The student was…

   sitting down. studying.

3. In her bathroom at home,^ Sally cleaned^ a/the sink^ with a wooden-handled^ bristle brush.^

   In the Campus Center women's room,^ Sally cleaned^ a/the sink^ with a wooden-handled^ bristle brush.^

   What was Sally doing with the sink?

   cleaning it using it

4. At his granddaughter's birthday party,^ Chuck admired^ a/the cake^ and let people know^ that he liked it.^

   In the fancy bakery,^ Chuck admired^ a/the cake^ and let people know^ that he liked it.^

   What was Chuck admiring?

   a cake a pie

5. In her living room,^ Susan tried out^ a/the hi-def television^ but was not^ completely pleased.

   ^ In the electronics store,^ Susan tried out^ a/the hi-def television^ but was not^ completely pleased.^

   Susan was checking out

   a TV. a radio.

6. As part of preparing his unicycle,^ Billy checked^ a/the tire^ for a puncture^ or a weak spot.^

   As part of preparing his bicycle,^ Billy checked^ a/the tire^ for a puncture^ or a weak spot.^

   Billy was fixing his…

   cycle. car.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

7. In her basement,^ Becky heard a bad noise coming from^ a/the washing machine^ and started^ worrying.^

In the laundromat,^ Becky heard a bad noise coming from^ a/the washing machine^ and started^ worrying.^

What probably needed fixing?

a washer a dryer

8. Sitting in front of his PC,^ Sam typed his name on^ a/the keyboard^ and listened to^ the clicks.^

Standing in the aisle at Staples,^ Sam typed his name on^ a/the keyboard^ and listened to^ the clicks.^

What was Sam checking out?

a computer part a TV part

9. Seeing his locker,^ Sandy tried to open^ a/the door^ but he simply^ couldn't manage.^

Seeing a row of lockers,^ Sandy tried to open^ a/the door^ but he simply^ couldn't manage.^

What did Sandy try to open?

a window a door

10. Picking up her hamburger,^ Patty noticed mold on^ a/the bun^ and was completely^ disgusted.^

Working the line at McDonalds,^ Patty noticed mold on^ a/the bun^ and was completely^ disgusted.^

What did Patty notice on the bun?

a bug mold

11. Playing near the pine tree,^ little Billy hid behind^ a/the trunk^ and was as quiet as^ he could be.^

In the grove of pine trees,^ little Billy hid behind^ a/the trunk^ and was as quiet as^ he could be.^

What did Billy hide behind?

a car a trunk

12. Standing in front of the Catholic church,^ Ann noticed some missing shingles on^ a/the steeple^ and felt sad.^

Standing in front of the row of churches,^ Annette noticed some missing shingles on^ a/the steeple^ and felt sad.^

What did Annette notice was missing?

some plants some shingles

13. Looking at the parked car,^ Fred thought he recognized^ a/the driver^ but realized^ he was mistaken about who it was.^

Looking at the line of cabs,^ Fred thought he recognized^ a/the driver^ but realized^ he was mistaken about who it was.^

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

What did Fred think he recognized?

a friend a driver

14. Putting the single flower into the vase,^ Abby had to trim^ a/the stem^ because the vase^ was too short to hold it.^

Putting the bunch of flowers into the vase,^ Abby had to trim^ a/the stem^ because the vase^ was too short to hold it.^

What did Abby put into the vase?

a plant a flower

15. Standing on his house porch,^ Joe put a package into^ a/the mailbox^ and walked^ back inside.^

Standing in his apartment complex lobby,^ Joe put a package into^ a/the mailbox^ and walked^ back inside.^

What was Joe doing?

picking up mail mailing something

16. Standing at his kitchen counter,^ Tom reached for^ a/the coffeepot^ and poured himself^ a big cup.^

Standing at the Dunkin Donuts self-serve counter,^ Tom reached for^ a/the coffeepot^ and poured himself^ a big cup.^

What did Tom reach for?

newspaper coffee pot

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

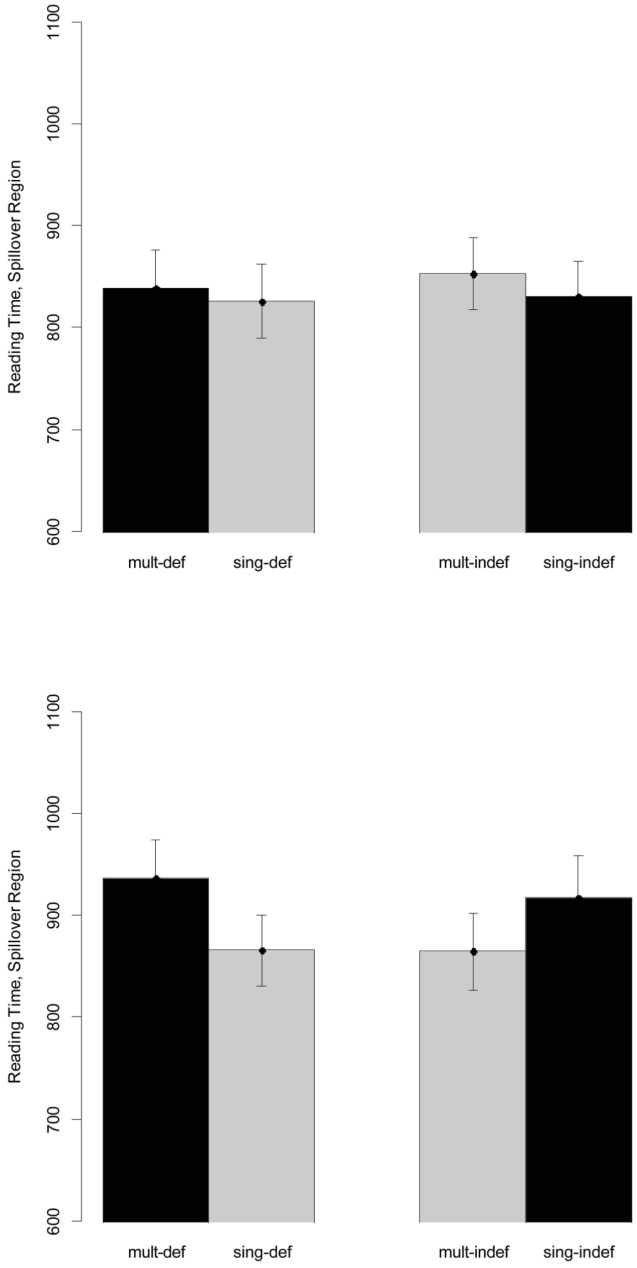

Figure 1.
Reading times (ms) for Region 4 (spillover), Experiment 1, no arithmetic task (top panel) and Experiment 2, arithmetic task (bottom panel), as a function of context and definiteness. Dark bars: Mismatch between context and definiteness; light bars: Match between context and definiteness. Standard error indicated.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

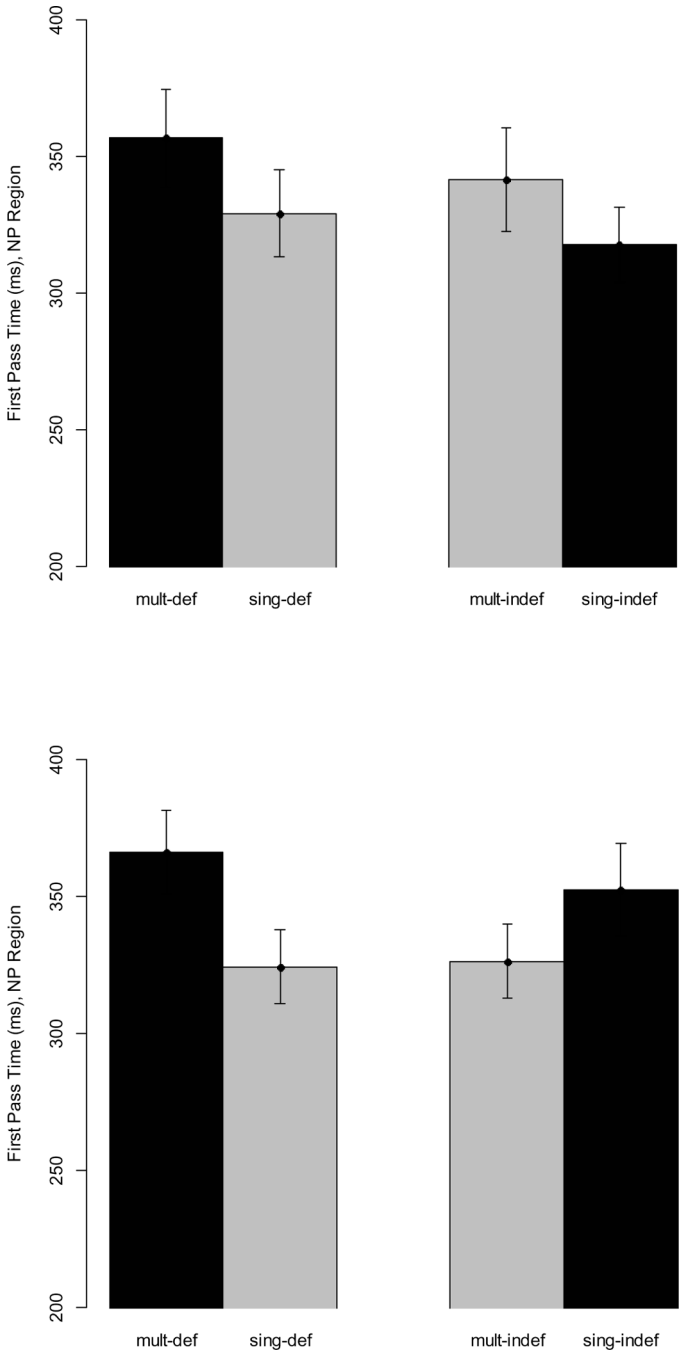

Figure 2.
First pass time (ms) for Region 3 (critical DP), Experiment 3, no arithmetic task (top panel) and Experiment 4, arithmetic task (bottom panel), as a function of context and definiteness. Dark bars: Mismatch between context and definiteness; light bars: Match between context and definiteness. Standard error indicated.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 1

Example of sentence set, Experiments 1 and 2

Condition	Sentence
Singular, indefinite	In the kitchen,^1 Jason checked out^2 a stove^3 very carefully.^4
Singular, definite	In the kitchen, ^ Jason checked out^ the stove^ very carefully. ^
Multiple, indefinite	In the appliance store, ^ Jason checked out^ a stove^ very carefully. ^
Multiple, definite	In the appliance store, ^ Jason checked out^ the stove^ very carefully. ^

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Mean reading times (ms) by region, for Experiment 1 (top panels) and Experiment 2 (bottom panels) (standard error in parentheses)

Experiment 1			
Condition	**Region 2**	**Region 3 (DP)**	**Region 4 (Spillover)**
Singular, indefinite	983 (45)	654 (23)	831 (35)
Singular, definite	1025 (49)	671 (25)	826 (36)
Multiple, indefinite	1084 (56)	599 (19)	853 (35)
Multiple, definite	1005 (49)	670 (23)	839 (37)

Experiment 2			
Condition	**Region 2**	**Region 3**	**Region 4**
Singular, indefinite	1077 (48)	655 (20)	917 (41)
Singular, definite	1081 (49)	693 (26)	866 (34)
Multiple, indefinite	1087 (45)	652 (23)	865 (38)
Multiple, definite	1136 (52)	689 (25)	937 (37)

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Example of sentence set, Experiments 3 and 4, with analysis regions indicated by ^

In the kitchen,^1 Jason checked out^2 a stove^3 very carefully^4 before using it.^5
In the kitchen,^ Jason checked out^ the stove^ very carefully^ before using it.^
In the appliance store,^ Jason checked out^ a stove^ very carefully^ before using it.^
In the appliance store,^ Jason checked out^ the stove^ very carefully^ before using it.^

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 4

Mean eyetracking (ms or proportions) by region, for Experiment 3 (standard error in parentheses)

Condition	Region				
	1	2	3 (DP)	4 (Spillover)	5
			First pass time		
Singular, indefinite	781 (36)	701 (35)	318 (14)	523 (25)	536 (29)
Singular, definite	786 (34)	742 (37)	329 (16)	509 (23)	578 (32)
Multiple, indefinite	886 (33)	751 (36)	341 (19)	511 (23)	568 (35)
Multiple, definite	886 (34)	742 (34)	357 (18)	455 (19)	552 (32)
			Regressions out (proportion)		
Singular, indefinite		.05 (.02)	.17 (.03)	.11 (.03)	.45 (.04)
Singular, definite		.03 (.01)	.13 (.03)	.07 (.02)	.40 (.05)
Multiple, indefinite		.05 (.02)	.16 (.03)	.12 (.02)	.46 (.05)
Multiple, definite		.03 (.01)	.07 (.02)	.12 (.03)	.45 (.05)
			Number of Fixations		
Singular, indefinite	3.78 (.17)	3.20 (.16)	1.29 (.06)	2.27 (.12)	2.25 (.14)
Singular, definite	3.88 (.16)	3.39 (.17)	1.38 (.07)	2.32 (.10)	2.01 (.12)
Multiple, indefinite	4.24 (.17)	3.34(.16)	1.28(.07)	2.12 (.10)	2.03 (.15)
Multiple, definite	.4.46 (.18)	3.28 (.15)	1.52 (.08)	2.02 (.08)	2.11 (.14)
			Second Pass time		
Singular, indefinite	82 (19)	119 (20)	79 (17)	142 (22)	90 (18)
Singular, definite	130 (29)	156 (35)	63 (17)	151 (22)	99 (22)
Multiple, indefinite	98 (21)	191 (32)	87 (15)	183 (24)	130 (30)
Multiple, definite	78 (23)	120 (30)	105 (25)	164 (25)	89 (21)

J Exp Psychol Learn Mem Cogn. Author manuscript; available in PMC 2014 March 01.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 5

Mean eyetracking (ms or proportions) by region, for Experiment 4 (standard error in parentheses)

Condition	Region				
	1	2	3 (DP)	4 (Spillover)	5
			First pass time		
Singular, indefinite	862 (37)	696 (34)	352 (17)	447 (21)	536 (30)
Singular, definite	878 (36)	747 (33)	324 (14)	427 (18)	583 (30)
Multiple, indefinite	1047 (37)	709 (29)	326 (13)	455 (22)	560 (29)
Multiple, definite	1077 (30)	705 (30)	366 (15)	466 (19)	552 (26)
			Regressions out (proportion)		
Singular, indefinite		.12 (.03)	.20 (.03)	.11 (.03)	.39 (.04)
Singular, definite		.07 (.02)	.16 (.03)	.09 (.02)	.42 (.04)
Multiple, indefinite		.06 (.02)	.17 (.03)	.08 (.02)	.45 (.04)
Multiple, definite		.05 (.02)	.12 (.03)	.10 (.02)	.40 (.04)
			Number of Fixations		
Singular, indefinite	4.38 (.19)	3.28 (.14)	1.51 (.07)	2.04 (.10)	2.14 (.13)
Singular, definite	4.65 (.20)	3.37 (.14)	1.47 (.07)	1.97 (.08)	2.34 (.13)
Multiple, indefinite	5.77 (.23)	3.29(.14)	1.38(.06)	2.04 (.09)	2.28 (.12)
Multiple, definite	5.56 (.18)	3.20 (.14)	1.62 (.07)	2.15 (.08)	2.18 (.12)
			Second Pass Time		
Singular, indefinite	253 (41)	300 (46)	136 (25)	143 (22)	115 (23)
Singular, definite	273 (40)	261 (35)	114 (19)	193 (27)	125 (22)
Multiple, indefinite	270 (43)	244 (31)	128 (18)	159 (20)	131 (23)
Multiple, definite	267 (42)	166 (29)	84 (13)	149 (21)	119 (21)

J Exp Psychol Learn Mem Cogn. Author manuscript; available in PMC 2014 March 01.