Published in final edited form as:
J Mem Lang. 2010 August 1; 63(2): 158–179. doi:10.1016/j.jml.2010.03.008.

Quantifiers more or less quantify online: ERP evidence for partial incremental interpretation

Thomas P. Urbach[1] and **Marta Kutas**[1,2]

[1]Department of Cognitive Science University of California, San Diego

[2]Department of Neurosciences University of California, San Diego

Abstract

Event-related brain potentials were recorded during RSVP reading to test the hypothesis that quantifier expressions are incrementally interpreted fully and immediately. In sentences tapping general knowledge (*Farmers grow crops/worms as their primary source of income*), Experiment 1 found larger N400s for atypical (*worms*) than typical objects (*crops*). Experiment 2 crossed object typicality with non-logical subject-noun phrase quantifiers (*most*, *few*). Off-line plausibility ratings exhibited the crossover interaction predicted by full quantifier interpretation: *Most farmers grow crops* and *Few farmers grow worms* were rated more plausible than *Most farmers grow worms* and *Few farmers grow crops*. Object N400s, although modulated in the expected direction, did not reverse. Experiment 3 replicated these findings with adverbial quantifiers (*Farmers often/rarely grow crops/worms*). Interpretation of quantifier expressions thus is neither fully immediate nor fully delayed. Furthermore, object atypicality was associated with a frontal slow positivity in few-type/rarely quantifier contexts, suggesting systematic processing differences among quantifier types.

Keywords

quantifier; incremental interpretation; brain potential; ERP; N400; language comprehension

1. INTRODUCTION

It is often important to specify amounts or quantities when communicating about objects and events. The number words in, *three balls and two strikes*, uttered during a baseball game provide quantitative information of critical importance to the parties involved. Natural languages have many ways to express quantity including grammatical determiners broadly construed, e.g., one, two, all, every, some, most, many, a few, nearly all, more than half, that modify nominal expressions, e.g., *outs*, *runners on base*, *pitchers* (Barwise & Cooper, 1981; Keenan & Stavi, 1986) and adverbs of quantification (Lewis, 1975), e.g., *often* and *rarely* in sentences like, *Batters rarely bunt with two strikes*, where they express information about the quantity or frequency of occurrences of events or event-like entities. It is uncontroversial

Corresponding author: Thomas P. Urbach, Department of Cognitive Science, Mail Code 0515, University of California, San Diego, La Jolla, CA, 92093-0515. Telephone 1-858-534-2440, FAX 1-858-534-1128, turbach@ucsd.edu.

that quantifier expressions systematically contribute to the overall meaning of the phrases and sentences in which they occur: *two outs with one runner on base* describes one sort of situation, *one out with two runners on base* describes quite another. However, the time-course of quantifier interpretation in real-time comprehension remains poorly understood. We conducted three rapid serial visual presentation (RSVP) reading experiments using event-related brain potentials (ERPs) to investigate *when* (immediately vs. delayed) and *to what extent* (fully vs. partially), the information afforded by simple quantifier expressions is integrated with world knowledge and incorporated into message-level representations during sentence comprehension.

1.1 Incremental interpretation and world knowledge

Sentence comprehension is rapid — skilled young adults can read for comprehension at rates of around 4-5 words per second (Just & Carpenter, 1980; Rayner, 1978). Comprehension is also generally thought to be incremental, i.e., lexical information is processed when a word is first encountered and then rapidly integrated with ongoing message-level representations at latencies on the order of hundreds of milliseconds. Incremental processing contrasts with a wait-and-see processing strategy on which multiple words may be buffered with interpretation delayed or deferred until other, perhaps critically informative words are encountered, e.g., at a clause or sentence boundary, with lexical and structural representations determined after what may be a substantial delay perhaps on the order of seconds (for recent overviews of incremental comprehension from different perspectives see Altmann & Mirkovic, 2009; Hagoort & van Berkum, 2007; Rayner & Clifton, 2009). A special case in the broader debate about incremental interpretation concerns the role of real-world or background knowledge (among the many other relevant factors). The details of how and when background knowledge constrains real-time comprehension are not fully understood although there is evidence from on-line measures that it can be brought to bear very rapidly. For example, Hagoort and colleagues recruited the N400 ERP to investigate the time course of the contribution of factual knowledge and conceptual knowledge to incremental comprehension (Hagoort, Hald, Bastiaansen, & Petersson, 2004). The N400 is a large (~ 5 μV) negative-going waveform typically beginning around 200ms and peaking around 400ms poststimulus (Kutas & Hillyard, 1980). The N400 is elicited by a variety of potentially meaningful stimuli including written and spoken words as well as pictures. N400 amplitude has been found to vary with a range of stimulus properties such as the frequency and concreteness and number of orthographic neighbors of the eliciting lexical item and is sensitive to a wide range of contextual factors involving aspects of word meaning, sentence meaning, and discourse context (for a review see Kutas, Van Petten, & Kluender, 2006). Perhaps the best-known finding is that words that are a poor semantic fit in context elicit a larger N400 than suitable control words, e.g., *Sue got up early and walked her* [*jet* / *dog*], though the more general finding is that larger N400 amplitudes are associated with words that are unexpected in context (Kutas & Hillyard, 1984). Hagoort, et al. (2004) noted that in Holland the trains are yellow and crowded, facts generally known to the Dutch, so for two sentences like, *The Dutch trains are* [*white* / *sour*] *and crowded*, Dutch people will know that both are false but for different reasons. The first is false because attempting to integrate the word *white* into the evolving representation of the sentence as an attribute of Dutch trains involves a failed correspondence with a well-known empirical fact. The second is false because attempting to integrate the word *sour*, an attribute of edible things, into the representation of the sentence as an attribute of the (inedible) Dutch trains involves a semantic feature mismatch. Hagoort and colleagues reasoned that if background knowledge of trains and semantic knowledge of word meanings contribute to comprehension in different ways or at different times, processing the semantically anomalous word, *sour* should differ from the factually incorrect word, *white*. They found, however, that both sentences elicited a large N400 in comparison

with the word, *yellow* in the true sentence, and, crucially, the N400 waveforms for the critical word in both false sentences did not differ in amplitude or latency. They interpreted this as evidence that background knowledge and lexical semantic information are integrated into the evolving interpretation on the same time scale and rapidly, i.e., within about 300ms. It is has not gone unnoticed that this argument is based on the failure to detect a difference and the question of whether background information is deployed as quickly as other types of information, e.g., lexical or conceptual information stored in semantic memory, remains somewhat controversial. Not withstanding temporally fine grained questions, on-line measures such as eye-movements (e.g., Ferguson & Sanford, 2008; Filik, 2008; Rayner, Warren, Juhasz, & Liversedge, 2004; Warren & McConnell, 2007; Warren, McConnell, & Rayner, 2008) and ERPs (e.g., Ferguson, Sanford, & Leuthold, 2008; Hagoort, et al., 2004; Nieuwland & Kuperberg, 2008; Nieuwland & Van Berkum, 2006) make a strong case that background knowledge is rapidly activated and deployed incrementally during comprehension.

1.2 Semantic Underspecification

At the same time, there is a cross-current to strong hypotheses about incremental interpretation, supported by a growing inventory of phenomena indicating that comprehenders may not fully process all the semantic information afforded by the verbal input and that the resulting message-level representations may be "partial" (Frazier & Rayner, 1990), "shallow" (Barton & Sanford, 1993), "underspecified" (Sanford & Sturt, 2002), or "good enough" (Ferreira, Bailey, & Ferraro, 2002). So-called semantic illusions, i.e., failures to detect false or semantically anomalous information, are a touchstone phenomenon, and may be observed in the lab by asking questions like, "How many animals of each type did Moses take on the ark?" (Erickson & Mattson, 1981) or "What is the holiday where children go door to door, dressed in costumes, giving out candy?" (Reder & Kusbit, 1991). The key findings are that people often fail to notice that Moses did not take the animals at all (it was Noah), and there is no such holiday (although on Halloween children often get candy). Other paradigms provide additional evidence that shallow semantic processing may be more widespread than first supposed. Frazier & Rayner (1990) used eye-movement data to argue that different meanings of lexically ambiguous words, e.g., *bank*, the financial institution vs. *bank*, the side of a river, are resolved immediately whereas sense differences, e.g., *newspaper* as the paper product in the driveway vs. the institution with an editorial policy are not. In their account, the representation of *newspaper* is initially underspecified with sense selection deferred until it becomes relevant for interpretation. Christianson, Hollingworth, Halliwell, & Ferreira (2001) found that after reading temporarily ambiguous garden-path sentences, e.g., *While Anna dressed the baby played in the crib*, people often responded "Yes" to the question, *Did Anna dress the baby*, even though this interpretation of the agent-action-patient thematic roles is inconsistent with globally correct syntactic structure. Sturt and colleagues (Sturt, Sanford, Stewart, & Dawydiak, 2004; Ward & Sturt, 2007) used a text-change detection paradigm to show that placing a critical entity in discourse focus, e.g., the word, *man*, in, *which man got into trouble* vs. *what was going on*, resulted in more detections of semantically close substitutions (… *the man in the* [hat / *cap*]) in a subsequent repetition of the passage. They argue that discourse focus leads to more detailed representations of the discourse entities that in turn allow fine-grained differences between *hat* and *cap* to be more readily noticed. The construction of these sorts of partially interpreted or semantically underspecified representations runs counter to a strong (immediate and full) incremental processing hypothesis.

1.3 Incremental quantifier interpretation

So, on the one hand, there is evidence that words are processed immediately and fully and that lexical semantic information and background knowledge are rapidly activated and integrated into evolving message level representations which, in turn, constrain the processing of subsequent words. On the other hand, there is also evidence that some semantic information such as the difference between *give* and *get* (Reder & Kusbit, 1991) may not be represented in the semantic context at all, let alone incrementally. What about quantifier expressions? On a strong incremental interpretation hypothesis, semantic information about quantity provided by the quantifier expressions in noun phrases such as, *two strikes*, *three ships*, and *most farmers*, is fully represented in the incrementally computed semantic representation of the noun phrase. If quantifier expressions are just another source of information that is fully and immediately incorporated into the evolving representation of semantic context they should have familiar sorts of processing consequences, e.g., constrain expectancies for upcoming information and facilitate or inhibit the access of information in semantic memory and its post-access integration into the current semantic representation.

Although the real-time processing of quantifiers has not been widely investigated, a number of special cases have been studied using online measures such as eye-movements, self-paced reading, and ERPs including the resolution of scope ambiguities in sentences containing multiple quantifiers, e.g., *Every kid climbs a tree* (Filik, Paterson, & Liversedge, 2004; Kurtzman & Macdonald, 1993; Paterson, Filik, & Liversedge, 2008), the role of quantifier expressions in modulating discourse focus (Moxey, Filik, & Paterson, 2009; Moxey, Sanford, & Dawydiak, 2001; Sanford, Dawydiak, & Moxey, 2007), and the resolution of ambiguous reference for bare cardinal quantifiers, e.g., *three ships* in simple discourse contexts such as, *Five ships sank. Three ships* … (Frazier, et al., 2005; Kaan, Dallas, & Barkley, 2007; Wijnen & Kaan, 2006).

Experimental evidence regarding the immediacy and depth of quantifier interpretation is mixed. Recent research suggests that cardinal determiners such as *three* are incrementally interpreted in discourse contexts where a set of ships has already been introduced, e.g., *Five ships appeared on the horizon* (Frazier, et al., 2005; Kaan, et al., 2007; Wijnen & Kaan, 2006). In such contexts, the determiner *three* that begins a subsequent sentence may end up serving different referential functions. If the noun phrase (NP) continues with a different noun, e.g. *planes*, it will introduce new entities into the discourse representation. However, if the NP is *three ships*, it will be ambiguous between picking out a subset of the given ships, i.e., three of the five just introduced, or introducing three additional ships into the discourse representation. Offline measures show that readers tend to preferentially resolve the referential ambiguity in favor of the subset interpretation (Frazier, et al., 2005; Wijnen & Kaan, 2006) and evidence from eye-movements (Frazier, et al., 2005), incremental behavioral measures (Wijnen & Kaan, 2006), and ERPs (Kaan, et al., 2007) indicates that these preferences are at work during on-line comprehension as well. For instance Frazier, et al. (2005 Experiment 2) recorded eye-movements while people read sentence pairs such as, Five ships appeared on the horizon. Three ships sank. In the critical comparison, this context was followed by, [*Two* / *Six*] *were bombarded by enemy fire*. They reasoned that if there was an on-line preference for ambiguous cardinal determiners to be assigned the subset interpretation, then there would be a processing disruption following the determiner, *Six*, because its cardinality precludes this interpretation. Consistent with the prediction, analysis of the eye-movements in the region immediately following the determiner (*were bombarded*) found increased first pass and total reading times following, *Six*, in comparison with, *Two*. This first-pass reading time effect is evidence that on-line comprehension processes register differences between these quantifier expressions when they are initially encountered and, furthermore, the direction of the effect (disruption for the interpretation

that is dispreferred off-line) is consistent with the idea that the initial on-line interpretation parallels the preferred offline resolution of the ambiguity. Further evidence comes from a related RSVP-ERP reading study (Kaan, et al., 2007). In this experiment, a short sentence introduced different numbers of entities into the discourse, e.g., [*Twelve* / *Four*] *flowers were put into the vase.* Both were followed by the same sentence, e.g., *Six had broken stems and were put in the trash*, in which the referentially ambiguous sentence initial cardinal determiner, *Six*, was compatible with the subset interpretation in the first context, i.e., six of the twelve flowers, but not in the second. Analysis of ERPs elicited by the critical word for all participants found *no* reliable effects before 900ms. Beginning around 900ms, a reliable broadly distributed relative positivity was observed when the determiner was incompatible with the subset interpretation preferred offline. This slow wave ERP effect emerges about half a second later than the first-pass reading time effects (Frazier, et al., 2005) though in the same two word region immediately following the critical determiner, e.g., *had broken*. This ERP effect is further evidence that the semantics of the quantifier expression is registered relatively rapidly: if not immediately, the delay is on the time-scale of words, not entire clauses.

These experiments with bare cardinal quantifiers provide evidence of incremental interpretation without violations of strong syntactic processing principles or semantic constraints. So, it would be natural to suppose that in sentences where quantifier interpretation leads to more salient semantic difficulties, evidence of on-line processing disruptions would be more pronounced and, perhaps, emerge more quickly. However the clearest direct empirical test we are aware of found precisely the opposite. In this ERP study (Kounios & Holcomb, 1992), nouns denoting categories and exemplars were combined with three quantifiers (*all*, *some*, *no*) in simple subject-predicate sentences presented in a speeded sentence verification task. Truth and falsity with respect to world knowledge were manipulated for both hierarchical category relations, e.g., *gems* (category) and *rubies* (exemplar), and relations between exemplars, e.g., *rubies* and *spruces*, in sentences such as the following (with nominal truth-value in parentheses): [*All* / *some* / *no gems*] *are spruces.* (F/F/T); [*All* / *some* / *no*] *spruces are gems.* (F/F/T); [*All* / *some* / *no*] *gems are rubies.* (F/T/F); [*All* / *some* / *no*] *rubies are gems*. (T/T/F). This experiment was not designed to investigate incremental interpretation per se and the stimuli were presented in two parts, the first consisting of the subject and copula, e.g., *All rubies are*, for 500ms, followed by a blank screen for 300ms, and then the predicate word, e.g., *gems* while ERPs were recorded. If determiners are interpreted incrementally, as suggested by the bare cardinal quantifier studies, then by the same reasoning Hagoort, et al. (2004) used in their investigation of the integration of word meaning and background knowledge, it might be predicted that when categorical background knowledge is activated by the quantified noun phrase, e.g., *All rubies are*, the final word, *gems*, when consistent with it should be relatively easier to process than when it is not, as in, *No rubies are gems*. However, based on an earlier finding (Fischler, Bloom, Childers, Roucos, & Perry, 1983) that N400 amplitude for the object noun phrase in sentences such as *A robin* [*is* / *is not*] *a bird*, did not vary with differences in truth-value resulting from the intervening *not*, Kounios and Holcomb predicted that their N400s would not be sensitive to differences in truth-value resulting from differences in the determiner. And, they were not. The sentence final N400 clearly reflected word-level semantic relations: *rubies* and *gems* had smaller N400s in the context of *gems* and *rubies* in comparison with *spruces* in these same contexts. More surprisingly perhaps, their manipulation of the determiner, e.g., *All rubies are gems* vs. *No rubies are gems* had no effect on the N400 elicited by *gems*. The authors' interpretation was that N400 reflects processing of semantic properties of words, e.g., categorical and associative relations between the nouns, but not the propositional and/or decision-making processes involved in working out the structural relations in the sentence or verifying the truth-value of the proposition expressed. The extent to which the findings generalize to other sentential

stimuli, presentation modes, and tasks is an open question. We also note that this null result, i.e., no N400 effect of determiner, may reflect a lack of power and/or sensitivity of the N400 with respect to those processes that vary as a function of the determiner semantics. However, even with these caveats, there is a *prima facie* dissociation between the way in which semantic information afforded by the determiner is processed and the way the semantic information afforded by the subject and predicate noun is processed. Since participant's truth-value judgments were generally accurate, it is clear that both quantifier and noun semantics were available to the system by the time these judgments were made. Yet even though the N400 is often sensitive to subtle manipulations of semantic context, there was no clear evidence in this case that the initial interpretation of the subject noun phrase and verb contains information that differentiates the determiners, *All* / *Some* / *No*.

The hypothesis that quantifier sentences are interpreted in stages with “logical” relations such as quantification and negation processed after subject-predicate relations is not new, (c.f., Carpenter & Just, 1975; Clark & Chase, 1972; Trabasso, Rollins, & Shaughnessy, 1971). Admittedly, however, this line of thinking has received little attention in an era where incremental interpretation is the received view. The emerging literature on semantic underspecification in language comprehension challenges strong formulations of incremental interpretation. And, although the Kounious & Holcomb (1992) ERP results concerning quantifier interpretation are not typically cited as examples of shallow interpretation, they do appear to pull in the same direction.

1.4 The present studies

A strong incremental interpretation hypothesis on which quantifiers are fully interpreted when initially encountered is consistent with the evidence from the bare cardinal experiments but is less obviously compatible with the Kounious & Holcomb (1992) findings. Since questions about the time course of quantifier interpretation remain, we conducted three experiments to investigate when (immediately vs. delayed) and to what extent (fully vs. partially) the semantic information afforded by two types of unambiguous quantifier expressions is processed. To that end, we pitted quantifier semantics, e.g., the meaning of *Few* and *Most* against background knowledge, e.g., of farmers and what they do.

In Experiment 1 we tested a baseline condition in which a bare plural subject noun and verb tap background knowledge and the typicality of the critical object noun varies, e.g., *Farmers grow crops* vs. *Farmers grow worms*. It is widely assumed that bare plurals involve an implicit generalization (for an overview see, e.g., Diesing, 1992). Full immediate incremental interpretation predicts that activation of the relevant background knowledge, associated relations, etc., in conjunction with implicit generalization will make *crops* easier to process than *worms* in this context, resulting in reduced N400 amplitudes for *crops* in comparison with *worms.* In Experiment 2 we pitted background knowledge of these same typical and atypical agent-action-patient contingencies against the meaning of explicit non-logical quantifier expressions in the subject noun phrase, e.g., [*Most* / *Few*] *farmers grow* [*crops* / *worms*]. In Experiment 3 we interposed adverbs of quantification between the bare plural subject noun and verb, e.g., *Farmers* [*often* / *rarely*] *grow* [*crops* / *worms*]. The semantics of these determiners and adverbs of quantification is either consistent with the background knowledge represented by the agent-action-patient combinations (Most farmers grow crops, Few farmers grow worms, Farmers often grow crops, Farmers rarely grow worms) or inconsistent with it (Few farmers grow crops, Most farmers grow worms, Farmers often grow worms, Farmers rarely grow crops). If the quantifiers are interpreted fully (vs. partially), these manipulations of the quantifier are predicted to reverse offline normative judgments that evaluate the proposition expressed against what is known, *e.g.*, *Most farmers grow crops* should be more plausible than *Most farmers grow worms*, and, this pattern should reverse for *Few farmers grow crops* and *Few farmers grow worms.*

Our primary interest concerns the time course of quantifier interpretation. According to strong formulations of incremental interpretation, c.f. the "immediacy assumption" (Just & Carpenter, 1980) and the "immediate complete interpretation" hypothesis articulated although not endorsed in Frazier & Rayner (1990), the semantic information afforded by the quantifier expression should be fully processed and integrated into the semantic and discourse context immediately as each word is encountered. On this hypothesis, initial processing of the critical object noun should be (relatively) facilitated when its typicality is consistent with the quantifier semantics in conjunction with background knowledge and (relatively) disrupted when it is not. With N400 amplitude as the online measure of processing difficulty, the full immediate quantifier interpretation hypothesis makes three specific predictions: 1. N400 amplitude for the typical object noun will vary as a function of the determiner with smaller N400 amplitude for *crops* in, *Most farmers grow crops* relative to *Few farmers grow crops*; 2. for the atypical object noun, the direction of this effect is predicted to reverse, with smaller N400 amplitude for *worms* in *Few farmers grow worms* relative to *Most farmers grow worms*; and, 3. the crucial prediction is that the N400 ERP typicality effect for *worms* vs. *crops* will reverse in the context of *Few farmers grow*, i.e., there will be a crossover interaction between quantifier and typicality for the online N400 amplitude effect that parallels the predicted crossover interaction in the offline plausibility judgments. The predictions are the same for Experiment 3 where the adverbs of quantification *often* and *rarely* are used in place of the subject noun phrase determiners such as *Most* and *Few*. In addition to testing the hypothesis for a lexically and structurally different type of quantifier expression, by reducing the number of words and, hence, the available processing time, between the quantifier expression and the critical test position at object noun, Experiment 3 provides an opportunity to replicate and extend the results of Experiment 2 and to sharpen inferences about the time course of incremental quantifier interpretation. All experiments reported below were conducted according to a research protocol approved by the Institutional Review Board of the University of California, San Diego Human Research Protection Program. Participants were volunteers who provided their informed consent in writing prior to enrolling in the study.

2. EXPERIMENT 1

2.1 Experiment 1 Methods

Participants—Thirty-two volunteers (mean age 21 years, range 18-37, 23 female) were recruited from the University of California, San Diego campus community. Volunteers received $7 per hour for participating and, at their discretion, could elect to apply 1 or 2 hours of participation toward course credit and receive $7 per hour for the balance of the time spent. All participants in these and subsequent experiments were right-handed, native English speakers with normal or corrected-to-normal vision and no reported history of neurocognitive impairment. Seven participants reported a left-handed parent or sibling. Data from one participant was excluded because of excessive EEG artifacts and an additional participant was recruited as a replacement.

Materials—Stimuli were constructed using 120 bare plural subject noun and transitive verb contexts denoting an agent and action, e.g., *Farmers grow.* Agents and actions were drawn from agent-action typicality norms (T. Ferretti, personal communication) with additional materials constructed by the experimenters. Each such context was paired with two object nouns, one denoting a typical patient, e.g., *crops*, and the other denoting an atypical object, e.g., *worms*, excepting one item where the atypical continuation was an adverb, *Joggers run* [*laps* / *monthly*] (see Table 1 for examples). None of the agent-patient contexts or object nouns were repeated and the typical and atypical object nouns were further constrained such that the mean length and frequency did not differ across the stimulus set: the mean log

Kucera-Francis frequency (Kucera & Francis, 1967) was 4.00 (*SD* = 2.99) for typical object nouns and 4.05 (*SD* = 2.89), $t(119) = 0.13$, $p = 0.895$; the mean character length was 5.50 (*SD* = 2.32) for typical object nouns and 5.33 (*SD* = 2.15), $t(119) = 0.75$, $p = 0.454$. After the object noun, the sentences continued with a phrase of between two and nine words long (median = 5, mode = 4) constructed to be semantically coherent with either object noun, e.g., *Farmers grow* [*crops* / *worms*] *as their primary source of cash*. These materials were combined with 90 sentences developed for an unrelated experiment that contained a variety of grammatical constructions, lengths, and degrees of contextual constraint. To avoid repetition with critical target words between the experiments, four pairs of the agent-action-patient sentences were excluded. The remaining 232 sentences were randomly assigned to two disjoint lists such that each list contained one member of each pair and a total of 58 atypical and 58 typical object nouns.

Procedure—Participants were seated in a comfortable chair in a dimly lit electrically shielded, sound attenuating testing chamber (Industrial Acoustics). Stimuli were presented under computer control on a 21″ VGA monitor in an amber colored font against a dark background at a viewing distance of about 120cm. Prior to the first word of the sentence, a fixation frame (~ 6 degrees of visual angle wide and ~ 2 degrees high) appeared and remained on while the sentence was presented word by word at an SOA of 500 ms, with each word appearing centered in the frame for a duration of 200ms. Stimuli were presented in blocks of 20 followed by a brief break. Following a random 25% of the sentences, a forced choice yes-no question appeared about 3s after offset of the final word that queried various aspects of the proposition expressed by the preceding sentence, e.g., *Did Charlie go to the park to fly a kite?* Participants indicated their answer via response buttons, (yes-no to left-right response hand mapping counterbalanced across subjects). Participants were instructed that they would be reading sentences one word at a time on the computer screen while their brainwaves were recorded and were told they would occasionally be asked to answer questions. They were encouraged to minimize eye-movements and blinks while the sentences were presented in order to reduce artifacts in the EEG. The instructions were followed by a brief practice session to familiarize participants with the stimulus presentation and task using sentences unrelated to the experimental materials.

EEG Data Recording and Analysis—Scalp ERPs were recorded from 26 electrodes embedded in an elastic cap as described in Ganis, Kutas & Sereno (1996), arrayed in a laterally symmetric quasi-geodesic pattern of triangles approximately 4 cm on a side (Figure 1, Panel A). An additional electrode was located over the right mastoid (A2); eye movements and blinks were monitored by recording the electro-oculogram (EOG) via four electrodes, one located adjacent to the outer canthus of and one below each eye. Potentials at all locations were recorded against a common reference electrode located over the left mastoid (A1), amplified with Grass Model 12 Neurodata Acquisition System (20K gain except for 10K gain at EOG and prefrontal locations, high pass filter 0.01 Hz, low pass filter 100 Hz), and digitally sampled (12-bits, 250 samples/s). Recordings were re-referenced offline to the mathematical average of the potentials at left and right mastoid. Single trial epochs spanning the interval from 500 ms prestimulus to 1500 ms poststimulus were extracted from the continuous EEG and screened for artifacts by computer algorithm and confirmed by visual inspection: 15% of the trials were excluded in each of the two conditions of experimental interest.

Time-domain average ERPs at the critical object noun position were computed for each participant. Mean amplitude relative to a 200 ms prestimulus baseline was computed for the object noun ERPs at the following latencies: P2 175-300 ms, N400 300-500 ms, late positivity (LP) 500-800 ms, and slow wave (SW) 800-1300 ms. Mean potentials were analyzed separately for the four midline electrodes and for sixteen of the remaining

electrodes at locations distributed across the scalp in a laterally symmetrical array (Figure 1, Panel A). For the midline electrodes we conducted a 2 × 4 repeated measures ANOVA with the stimulus factor of object noun typicality (typical, atypical) fully crossed with the electrode location factor of anteriority (Pf, Ce, Pa, Oc). For the 16 mediolateral electrodes we conducted a 2 × 2 × 2 × 4 ANOVA fully crossing typicality with electrode location factors of hemisphere (left, right), laterality (lateral, medial), and anteriority (prefrontal, frontal, temporo-central, parieto-occipital). For *F* tests involving more than one degree of freedom in the numerator, we report *p* values for Greenhouse-Geisser epsilon-adjusted degrees of freedom (Greenhouse & Geisser, 1959), the value of epsilon, and the original (unadjusted) degrees of freedom. ANOVAs were conducted using Cleave, an open source data analysis utility (Herron, 2005). Figures were constructed using open-source software (ggplot2, http://had.co.nz/ggplot2, Wickham, 2009; Inkscape, http://www.inkscape.org, Bah, 2007). Since no reliable P2 effects were observed in Experiment 1 or subsequent experiments, we omit the results of the P2 analyses.

2.2 Experiment 1 Results

The ERP morphology at the noun was typical for the 500ms SOA RSVP paradigm (Figure 1, Panel B). P1-N1-P2 potentials over lateral occipital scalp were observed between 50 and 200ms poststimulus followed by a large P2 over frontocentral scalp peaking shortly after 200ms. Following the P2, a large broadly distributed negative going deflection peaking about 400ms (N400) was observed in both conditions. The N400 waveforms elicited by atypical and typical object nouns begin to diverge about 300ms poststimulus onset at all but the prefrontal electrode locations, atypical more negative-going, and this difference reaches a maximum at about 400ms poststimulus. The main effect of typicality was reliable at midline and mediolateral electrodes (ANOVAs in Table 2). The effect is broadly distributed across the scalp (Figure 1, Panel C), largest at medial, centroparietal locations, and slightly right lateralized resulting in interactions between typicality and electrode location factors of laterality, anteriority, and hemisphere (ANOVAs in Table 2).

Following the N400 and superimposed upon the visual evoked potential wavetrain elicited by the subsequent word there is a smaller relative negativity for atypical in comparison with typical object nouns that persists throughout the balance of the epoch over medial scalp posterior to the prefrontal electrodes. At the midline electrodes, the effect reverses slightly at the prefrontal electrode where atypical nouns are more positive than typical nouns during the LP time window (500-800ms) and SW time window (800-1300ms). At mediolateral electrodes, the atypical nouns were slightly more positive at prefrontal electrodes in the LP time window and more negative at medial posterior electrodes, with this posterior negativity somewhat larger at right in comparison with left medial electrodes. Similar effects were observed in the SW time window except for the left-right asymmetry. These distributional differences resulted in interactions between typicality and electrode location factors in each time window (ANOVAs in Table 2).

2.3 Experiment 1 Discussion

Experiment 1 confirmed that the sentence context consisting of a bare plural subject noun and transitive verb already establish sufficient semantic context to modulate processing of the typical and atypical object nouns during word by word sentence reading. As predicted, the atypical object nouns in sentence contexts like, *Farmers grow wormsas their primary source of cash*, elicited a clear N400 effect in comparison with the typical object nouns, e.g. *crops*, in the same context. This N400 effect was unexceptional with respect to the latency, polarity, and scalp distribution and crucially, does not involve a semantic anomaly or incongruity. We interpret this N400 effect as evidence of a processing difference that depends upon what people know about farmers, crops, worms, and what farmers do, c.f.,

Dutch trains are [*white* / *yellow*] (Hagoort, et al., 2004). These results are consistent with the predictions of incremental processing models on which background knowledge about the denoted agent and action is rapidly activated and available to constrain the processing of subsequent words whether at the level of lexical access, post-access integration or both. From the direction of the N400 effect, we infer that processing of *crops* is relatively facilitated in comparison with *worms*, consistent with semantic models that treat bare plurals as implicit generalizations even in the absence of an overt quantifier expression. Although the contribution of lexical level processing, e.g., semantic priming of *crops* by *grow* cannot be dissociated from presumed sentence-level processes in this design, observing this N400 effect in the expected direction provides a key comparison with Experiment 2 and Experiment 3 where sentence level processing is manipulated to test hypotheses about the online processing of quantifier expressions. Furthermore, subsequent to the N400 we observed a sustained posterior negativity in conjunction with a small, generally prefrontal positivity. Strong conclusions about the functional significance of the prefrontal positivity cannot be drawn from this two-way comparison. These later effects may reflect a continuation of the processing associated with the amplitude modulation of the N400 or functionally distinct processing that occurs afterwards or both.

3. EXPERIMENT 2

Experiment 2 extends Experiment 1 by pitting the semantics of explicit quantifier expressions against background knowledge and comparing the consequences of the quantifier semantics for offline interpretation with their effects during incremental comprehension. The hypothesis that quantifiers are interpreted fully and immediately predicts that the offline interpretations are computed on-line. We tested this prediction by comparing the pattern of offline plausibility judgments with on-line N400 evidence of processing disruptions.

3.1 Experiment 2 Methods

Participants—A new group of 20 adult volunteers (mean age = 20 years, range 18 - 24, 10 women) were recruited from the University of California, San Diego community and participated for course credit or for cash. Eight participants reported a left-handed parent or sibling. EEG data from two participants was excluded because of excessive EEG artifacts and two additional participants were recruited as replacements.

Materials—The stimuli in Experiment 2 were constructed from those in Experiment 1 by preceding the bare plural subject nouns with a determiner to form a quantified subject noun phrase. The determiners were of two quantificational types, grouped according to whether they picked out relatively larger or smaller sets of objects, e.g., *Most farmers* and *Few farmers*, respectively. These determiners which we descriptively label “most-type” and “few-type” were matched for the number of words in the following eight pairs: Most/Few, Many/Few, Almost all/Almost no, Practically all/Practically no, A large number of/A small number of, Nearly all/Rather few, Lots of/Hardly any, A lot of/A very few. As a group, the few-type quantifiers are “negative” in the sense that they license negative polarity items, e.g., *ever* (Fauconnier, 1975; Krifka, 1995): compare the ill-formed most-type sentence, Many college baseball players *ever reach the pros, with the well-formed corresponding few-type sentences, [Few / Almost no / Practically no / Rather few / Hardly any] *college baseball players ever reach the pros*, though two cases may be less clear: [*A small number of / A very few*] *college baseball players ?ever reach the pros*. These eight pairs of most- and few-type determiners were distributed among the 120 sentence pairs in Experiment 1 to obtain 120 sets of sentences in four conditions (see Table 1 for examples). These sentences were assigned at random to four disjoint lists such that each list contained an equal number

of sentences with typical and atypical objects (60 each), an equal number of sentences with most-type and few-type quantifiers (60 each) and an equal number of the four combinations of determiners and object nouns (30 each). The determiner expressions were also distributed across the four lists such that each list contains four different most-type and four different-few type quantifier expressions and half of each type occurred with typical and atypical object nouns. Across lists, each member of the eight quantifier expression pairs appears equally often with typical and atypical objects. An oversight in the counterbalancing scheme resulted in a systematic relation between the occurrences of the quantifier expression, e.g., *Most*, or *Hardly any*, and the nominal typicality (though not identity) of the object noun. This relation was obscured by the variety of quantifier expressions, object nouns, and fillers and there was no evidence from debriefing that participants were aware of it. Since the results in this experiment are qualitatively and quantitatively similar to those in Experiment 3 where there was no such contingency, this relation seems unlikely to play a significant role in the results or conclusions. These 120 quantifier materials were combined with an additional 88 filler items of two sorts. Sixty were complex sentences containing a coordinating conjunction, of which half involved a verb-sense shift, e.g., Mounties hiked the fees and the trails at the park, and half did not, e.g., Mounties hiked the paths and the trails at the park. An additional 28 sentences were of a variety of grammatical forms and half ended with a final word that was possible but unlikely in context, e.g., *Amy woke early every morning to walk her* [*dog* / *cow*].

Procedure—Stimulus presentation was as described for Experiment 1 except that 1800 ms after each sentence, a prompt appeared, *How plausible*? And participants indicated their rating on a 5-point scale (1=highly implausible, 2=moderately implausible, 3=neutral, 4=moderately plausible, 5=highly plausible) by pressing one of five labeled buttons mounted on a panel with the thumb and four fingers of the right hand. Responses were not speeded though participants were encouraged to respond based on their initial impression. A card below the computer monitor displayed the plausibility-scale-to-response-button mapping throughout the experiment. Participant's response to the plausibility question was followed by a brief pause and then presentation of the next sentence.

Plausibility judgment analysis—Summary measures of offline plausibility were computed for each subject as the weighted average of their plausibility judgments in each condition. Although the responses were not speeded, they were timed and on grounds that exceptionally long response times may reflect the intrusion of qualitatively different processing, summary scores were also computed after excluding those responses with latencies greater than 3 times the interquartile range above the 3rd quartile for each subject (Tukey, 1977). These response time outliers comprised about 2.5% of the data. In separate analyses of the complete and trimmed sets of responses, none of the experimental effects differed in direction or statistical reliability and we report the results for the trimmed data. A 2-way repeated-measures ANOVA was conducted on the mean plausibility ratings with two levels of quantifier (most-type, few-type) and two levels of object noun (typical, atypical). Planned tests of the effects of quantifiers on plausibility judgments were conducted with paired-sample Welch *t* tests, two-tailed except for one-tailed tests of effects in a predicted direction in which case we report probabilities as $p_{1\text{-tailed}}$ (t.test function in R 2.9.0, R Development Core Team, 2009). For these *t* tests we report Cohen's paired-sample *d* (Cohen, 1988) as a measure of effect size and characterize effects as small, medium, and large at $d = 0.2$, 0.5, and 0.8 respectively.

EEG data recording and analysis—EEG data acquisition, screening, and ERP data reduction for potentials elicited by the critical object noun (N400, LP, SW) were all conducted as described for Experiment 1. In the conditions of experimental interest, on

average between 4% and 6% of the trials contained EEG artifacts and were excluded from subsequent analyses. ANOVAs were conducted as in Experiment 1 except for the addition of the within-subjects factor of quantifier (most-type, few-type) in the ANOVA and planned comparisons to test the predicted effects of quantifiers on N400 amplitude. For the midline electrodes we thus conducted a Quantifier × Typicality × Anteriority ANOVA and for the sixteen mediolateral electrodes we conducted a Quantifier × Typicality × Hemisphere × Laterality × Anteriority ANOVA. The predicted effects of quantifiers on N400 amplitude were tested in the same manner as the plausibility judgments, via paired-sample Welch *t* tests on mean amplitude 300 – 500ms poststimulus at midline and mediolateral electrode locations posterior to the prefrontal electrodes. Since the 200ms prestimulus baseline corresponds to the N400 of the previous word, we also measured and analyzed poststimulus potentials relative to a shorter (100ms) and longer (500ms) prestimulus baseline. The patterns of effects were similar regardless of the choice of baseline and we report results for the 200ms prestimulus baseline analysis.

3.2 Experiment 2 Results

Plausibility judgments—Both the quantifier and object noun manipulations had clear effects on the plausibility ratings. There was a robust main effect of quantifier type with sentences containing most-type quantifiers rated more plausible (mean = 3.3, *SD* = 1.15) than those few-type quantifiers (mean = 2.6, *SD* = 0.64), $F(1,19)=79.58$, $MSE=0.10$, $p < .001$, $\eta_p^2 = .81$. There was also a reliable main effect of typicality with sentences containing typical object nouns rated more plausible (mean = 3.27, *SD* = 1.14) than those containing atypical object nouns (mean = 2.6, *SD* = 0.65), $F(1,19) = 43.23$, $MSE = 0.18$, $p < .001$, $\eta_p = .69$. Crucially, the quantifier and object noun typicality factors exhibited the predicted crossover interaction, $F(1,19) = 248.06$, $MSE = 0.18$, $p < .001$, $\eta_p = .93$, (Figure 2, Panel A). As expected, sentences containing a typical object noun and beginning with a most-type quantifier were more plausible (mean = 4.3, *SD* = 0.29) than those beginning with a few-type quantifier (mean = 2.2, SD = 0.43), $t(19) = 22.09$, $p_{1\text{-tailed}} < 0.001$, $d = 4.92$. When the sentences contained an atypical noun this pattern reversed and sentences beginning with a few-type quantifier were reliably more plausible (mean = 3.1, *SD* = 0.49) than those beginning with a most-type quantifier (mean = 2.2, *SD* = 0.45), $t(19) = -6.486$, $p_{1\text{-tailed}} < 0.001$, $d = -1.45$. Critically, the few-type quantifiers did not merely modulate the off-line plausibility ratings but fully reversed them such that sentences with atypical objects were rated more plausible than those with typical objects, $t(19) = -6.513$, $p < 0.001$, $d = -1.45$. This crossover interaction in the off-line plausibility ratings is an important point of contrast with the ERP results.

ERPs—The ERP morphology in Experiment 2 was again typical for the 500ms SOA RSVP paradigm (Figure 2 Panel B). P1-N1-P2 potentials over lateral occipital scalp were observed between 50 and 200ms poststimulus followed by a large P2 over frontocentral scalp peaking shortly after 200ms. A large broadly distributed negative going deflection with an onset shortly before 300 ms and peaking about 400ms (N400) was observed in all conditions, with amplitude modulated by the experimental manipulation of object noun typicality and quantifier. From about 300-500 ms a broadly distributed relative negativity (N400 effect) is observed for the atypical object nouns in comparison with typical object nouns and the amplitude of this N400 effect (~ 2 – 3 μV) is modulated to a lesser degree (~ 1μV or less) by the quantifier type (Figure 2 Panel B, blue shading). Among the four experimental conditions, the maximum and minimum N400 amplitudes occur in the context of the most-type quantifiers in the expected direction: the largest (most negative) N400 is elicited by atypical object nouns in the context of most-type quantifiers, e.g., *Most farmers grow worms*, and the smallest by typical object nouns, e.g., *Most farmers grow crops*. The N400 amplitudes associated with object nouns in the context of the few-type quantifiers fall

between these extrema. Specifically, the N400 associated with the typical object nouns in the context of few-type quantifiers, e.g., *Few farmers grow crops*, is slightly larger in comparison with the N400 elicited by these same words in the context of the most-type quantifiers, e.g., *Most farmers grow crops*. N400 amplitude associated with atypical object nouns in the context of few-type quantifiers, e.g., *Few farmers grow worms*, is slightly smaller than the N400 for these same object nouns in the context of the most-type quantifiers, e.g. *Most farmers grow worms*. Both these modulations of N400 amplitude by the few-type quantifiers were in the expected direction, i.e., N400 increase for typical object nouns and N400 reduction for atypical object nouns but, crucially, the object noun typicality N400 effect does not reverse in the context of the few-type quantifiers (compare Figure 2, Panels A and C). Following the N400 effect, a late positive deflection overlapping the P2 of the following word is observed at prefrontal and frontal electrodes, largest following atypical object nouns in the context of few-type quantifiers, e.g., *Few farmers grow worms*, intermediate for both typical and atypical object nouns in the context of most-type quantifiers, e.g., *Most farmers grow* [*crops* / *worms*], and smallest for typical object nouns in the context of few-type quantifiers, e.g., *Few farmers grow crops* (Figure 2, Panel B, red shading).

Early potentials: Based on visual inspection, exploratory ANOVAs were conducted on mean amplitudes early in the epoch. In the 0 – 50ms poststimulus window, no effects involving the quantifier or typicality factors were reliable across all three baselines. For potentials in the 50 – 150ms window, typicality interacted with electrode location factors for the mediolateral electrodes only (ANOVAs in Table 3).

N400 (300-500ms): At the midline electrodes the main effect of object noun typicality for N400 amplitude accounted for a substantial amount of variability (ANOVAs in Table 3). The main effect of quantifier and interactions between quantifier and anteriority were not reliable. In the comparisons of primary interest, the few-type quantifiers reliably modulated N400 for the typical nouns in the expected direction, increasing the (negative-going) N400 amplitude for the typical object nouns from −0.36 μV (*SD* = 2.09 μV) in the context of the most-type quantifiers to −1.17 μV (*SD* = 1.37 μV), $t(19) = 2.03$, $p_{1\text{-tailed}} = 0.029$, $d = 0.45$. For the atypical nouns, the numerical decrease in N400 amplitude from −2.87 μV (*SD* = 1.794) μV in the context of most-type quantifiers to −2.50 μV (*SD* = 1.474) μV in the context of few-type was not reliable ($p_{1\text{-tailed}} > 0.88$). In the critical test of whether the few-type quantifiers would reverse the typicality effect, it is clear from inspection of the waveforms that there was no crossover effect in the context of the few-type quantifiers, and the N400 elicited by the atypical nouns remained reliably more negative than for the typical nouns, $t(19) = 3.53$, $p = 0.002$, $d = 0.78$. The pattern of effects at the mediolateral electrodes was qualitatively similar, with atypical nouns eliciting a larger N400 than typical. This effect was larger at medial in comparison with lateral electrodes, and larger over the right hemisphere at locations posterior to prefrontal electrodes (ANOVAs in Table 3). The main effect of quantifier and interactions between quantifier and electrode location factors were not reliable. For the planned comparisons of primary interest, quantifier effects on N400 amplitude at mediolateral electrodes were in the expected direction though smaller than at the midline locations, being marginal for the typical objects ($p_{1\text{-tailed}} = 0.065$) and again not reliable for atypical objects ($p_{1\text{-tailed}} > 0.84$). Again, there was no reversal of the typicality effect in the context of few-type quantifiers with N400 for the atypical objects (mean = −1.82 μV, *SD* =1.35 μV) remaining reliably more negative than for typical (mean = −0.61 μV, *SD* = 0.10 μV), $t(19) = 4.29$, $p < 0.001$, $d = 0.96$.

LPC (500-800ms): The LPC appears to be a transition between the end of the N400 typicality effect at the posterior electrodes and the beginning of the frontal positivity observed later in the epoch. At the midline electrodes, there were no reliable effects of

quantifier, typicality, or interactions between these factors and electrode locations. At the mediolateral electrodes, the main effect of quantifier was not reliable and this factor did not interact with any electrode location factors. In comparison with typical nouns, the potentials elicited by atypical nouns continued to be more negative at right posterior electrode locations and over the left hemisphere, the effect was smaller at medial locations and reversed polarity at left lateral frontal and central electrodes (ANOVAs in Table 3). The scalp distribution further varied as a function of quantifier and typicality, most saliently for the atypical nouns. In the context of the most-type quantifiers, atypical nouns are the most relatively negative of the four conditions over right frontal and central scalp and in the context of the few-type quantifiers they are the most positive at prefrontal electrodes (ANOVAs in Table 3). Absent a priori predictions about the distribution of these effects, we conducted pairwise *t* tests at each electrode location and determined that the largest effect was of medium size ($d = 0.57$) and occurred at the left medial prefrontal electrode where the LPC for atypical nouns was relatively more positive in the context of few-type quantifiers (3.33 μV) in comparison with most-type (1.85 μV).

Slow wave (800-1300ms): At the midline electrodes ANOVA found no reliable effects of quantifier, typicality, or interactions between these factors and electrode locations. At the mediolateral electrodes (Figure 2, Panel D, Slow Wave), potentials for both types of object nouns following the few-type quantifiers tended to be more positive at medial electrodes and lateral prefrontal electrodes, (ANOVAs in Table 3). Potentials elicited by atypical nouns were more positive than typical nouns at most locations and this effect was larger over medial than lateral prefrontal scalp, and slightly reversed at lateral occipital locations. The anterior positivity for atypical in comparison with typical nouns was greatest in the context of the few-type quantifiers, somewhat larger at left medial anterior electrodes, and this effect reversed polarity over central and posterior scalp with slightly greater relative negativity over right than left occipital electrodes. These distributional differences resulted in the interactions between the quantifier, typicality and electrode location factors summarized in Table 3. Salient among these effects is the prominent typicality effect in the context of few-type quantifiers (Figure 2, Panel B, red shading). We computed effect sizes at each electrode and found that largest effect sizes for this positivity were observed at medial prefrontal electrodes: left (1.94 μV, $d = 1.01$); midline (1.84 μV, $d = 0.91$); right (1.81 μV, $d = 0.73$). The effect sizes also tended to be somewhat larger over the left in comparison with homologous right prefrontal locations, and decreased from front to back with $|d| < 0.27$ at al locations posterior to frontal scalp. By contrast, in the context of the most-type quantifiers, the maximum object typicality effect size anywhere on the scalp was small ($d = 0.34$ at the right lateral occipital electrode) and smaller still at the prefrontal electrodes, $0.003 \leq |d| \leq 0.21$. In sum, this slow wave effect appears best described as an object noun typicality effect manifest as a predominantly medial prefrontal, slightly left lateralized positivity observed for atypical objects in the context of few- but not most-type quantifiers.

3.3 Experiment 2 Discussion

Experiment 2 investigated the incremental interpretation of quantified subject noun phrases. As expected, sentences beginning with most-type quantifiers and containing typical object nouns were rated more plausible than those with atypical object nouns. The few-type quantifiers provide the crucial test of the full quantifier interpretation hypothesis. If people assigned the quantifiers a full (as opposed to partial or underspecified) interpretation and the resulting message-level representation of the sentence is integrated with background knowledge of agent-action-patient contingencies in making plausibility judgments, few-type quantifiers should reverse the plausibility judgments regarding typical and atypical object nouns: they did. This crossover interaction was not entirely symmetric in that few-type quantifiers with the atypical object nouns, although reliably more plausible than with typical

object nouns, were nonetheless still less plausible than most-type quantifiers with typical object nouns. This may reflect background knowledge about the atypical object nouns selected for the comparison or a general bias against rating *few* type sentences toward the higher end of the plausibility scale.

With clear evidence about the end state of quantifier interpretation the key question is how this interpretation was computed in real time. The full immediate quantifier interpretation hypothesis leads us to consider the possibility that the interpretation of a quantified subject noun phase by itself establishes semantic contexts that, along with background knowledge, could differentially affect the processing of subsequent words. The analysis of ERPs at the typical and atypical object noun provide a sharp contrast with the offline plausibility judgments. Consistent with full incremental interpretation we found that atypical object nouns elicited relatively greater N400 amplitude than typical nouns in the context of most-type quantifiers, e.g., *Most farmers grow* [*crops* / *worms*]. By the same line of reasoning, full incremental interpretation of *Few* in the context of *Few farmers grow* should facilitate processing of words denoting things typically grown by few farmers, e.g., *worms*, in comparison with things not typically grown by few farmers, e.g., *crops* in which case, the N400 amplitude effect should be reversed. In the crucial test, we found that, contrary to this prediction, *worms*, still elicited a larger N400 than *crops* in sentences like, *Few farmers grow* [*crops* / *worms*], c.f. the larger N400 for *spruces* in, *No rubies are* [*gems* / *spruces*] (Kounios & Holcomb, 1992). So although the offline plausibility judgments for both types of quantifiers were consistent with full quantifier interpretation and immediate integration with background knowledge, the online N400 measures were not and failed, in particular, for the few-type quantifiers. An alternative to full immediate incremental interpretation is fully deferred quantifier processing wherein quantifier semantics are initially unspecified, with interpretation occurring later at a significant delay. On this view, the incremental interpretations of the two contexts, [*Most* / *Few*] *farmers grow* are initially identical, in which case there should be no differential facilitation of *crops* or *worms* when these words are first encountered and processing differences that depend on the specific meaning of the quantifier are predicted to evolve later. However, unlike Kounios & Holcomb (1992), we found that the different quantifiers reliably modulated midline N400 for typical (but not atypical) nouns. This quantifier effect on object noun processing is evidence that the incrementally computed semantic contexts for sentences with the different quantifier types are not identical. Since other contextual factors that may modulate N400 amplitude, e.g., lexical associations between *grow* and *crops*, and frequency of usage, are held constant in this experimental design, we attribute the N400 modulation to the experimental manipulation of the quantifier. Furthermore, the direction of the effect, and the fact that it is observed on the N400 are evidence that the effect is related to the appropriate meaning of the quantifier being incrementally incorporated into the evolving semantic context. Taken together, these ERP results—modulation of N400 amplitude by quantifiers in the expected direction but short of the crossover effect observed for plausibility judgments—argue against both the hypothesis that quantifier interpretation is full and immediate and the hypothesis that quantifier interpretation is fully delayed. These effects seem better explained by the hypothesis that quantifier interpretation is incremental but that these initial interpretations are partial or underspecified in comparison with the representations used in making subsequent plausibility judgments.

We also observed a prefrontal slow wave occurring after the N400 that exhibited a different pattern. For the few- but not most-type quantifiers, a prefrontal effect was observed in the later time windows with atypical nouns relatively more positive than typical. This finding indicates that the time course of processing the most- and few-type quantifiers differs and suggests that at least some aspect(s) of the processing of few-type quantifiers is delayed relative to most-type quantifiers. Although the functional significance of this prefrontal

positivity is not known, these results together already argue against any real-time processing hypothesis that does not allow for systematic differences between types of quantifiers.

4. Experiment 3

In Experiment 3 we aimed to replicate Experiment 2 and test our conclusions about incremental quantifier interpretation with regard to a different type of quantifier expression: adverbs of quantification, e.g., *Farmers* [*often / rarely*] *grow crops*. With only one word intervening between the quantifier expression and the critical object noun, these sentences afford the processer less time to activate and integrate the quantifier semantics and background knowledge than did the sentences in Experiment 1 with quantified subject noun phrases so these materials provide a stronger test of the hypothesis that the interpretation of expressions of quantity is incremental. The predictions for the [*often* / *rarely*] comparisons in Experiment 3 are analogous to those for the [*most* / *few*] comparisons of Experiment 2. We expected to find a Quantifier × Typicality crossover interaction effect in the plausibility ratings. Again, the key test of incremental quantifier interpretation was whether the processing difficulty associated with atypical relative to typical object nouns when first encountered is modulated by the quantifier manipulation and if so, whether it is reversed in parallel with the offline plausibility judgments.

4.1 Experiment 3 Methods

The methods for Experiment 3 were the same as described for Experiment 2 with the exception of the participants, stimulus materials, and the addition of planned comparisons at selected electrode locations in the LPC and SW time windows based on the effects observed in Experiment 2. A new group of 24 adult volunteers (mean age = 20, range 18–22, 13 women) were recruited from the University of California, San Diego community and participated for course credit or for cash. Eight participants reported a left-handed parent or sibling. EEG data from four participants were excluded because of excessive EEG artifacts and four additional participants were recruited as replacements. The stimuli for Experiment 3 were again constructed from the stimuli used in Experiment 1, this time by interposing one of two adverbs of quantification, *often* or *rarely*, between the bare plural subject noun (see Table 1 for examples). These sentences were assigned to four lists such that one sentence from each of the 120 sets appeared on each list and each list contained 60 sentences in each of the four experimental conditions. These lists were combined with the same filler items described for Experiment 2. Stimulus presentation, behavioral and EEG data acquisition and analysis procedures were the same as in Experiment 2 with the addition of planned comparisons for the prefrontal quantifier and typicality effects in the LPC and SW latency windows based on the effects observed in Experiment 2. In four participants eye-blinks (20-40% of the trials, mean 28%) were corrected using an adaptive spatial filter (Dale, 1994) and after blink correction on average between 9% and 10% of the trials in conditions of experimental interest were excluded from subsequent analysis. Plausibility rating response time outliers were defined as in Experiment 2 and trimming excluded 2.8% of the responses from the analysis.

4.2 Experiment 3 Results

Plausibility Judgments—Both the adverb of quantification and object noun manipulations again had clear effects on the plausibility ratings and interacted in the predicted direction (Figure 3, Panel A). There was a numerically small but statistically robust main effect of quantifier type with sentences containing, *often*, rated more plausible on average (mean = 3.1, *SD* = 1.36) than those containing, *rarely*, (mean = 2.8, *SD* = 0.52), $F(1,23) = 21.57$, $MSE = 0.08$, $p < .001$, $\eta_p^2 = .48$. There was also a main effect of typicality with sentences containing typical object nouns rated more plausible (mean = 3.46, *SD* =

1.04) than those containing atypical object nouns (mean = 2.49, *SD* = 0.77), $F(1,23) = 271.74$, *MSE* = 0.08, $p < .001$, $\eta_p^2 = .92$. As predicted, there was a reliable crossover interaction effect $F(1,23) = 238.44$, *MSE* = 0.27, $p < .001$, $\eta_p^2 = .91$. Planned comparisons found that, as expected, for the typical objects, sentences containing the adverb *often* were reliably more plausible (mean = 4.4, *SD* = 0.33) than those containing *rarely*, (mean = 2.51, *SD* = 0.45), $t(23) = 14.81$, $p_{1\text{-tailed}} < 0.001$, $d = 3.02$. Also as expected, for the atypical objects, sentences with the adverb *rarely* (mean = 3.17, *SD* = 0.36) were reliably more plausible than those with *often* (mean = 1.80, *SD* = 0.32), $t(23) = -12.23$, $p_{1\text{-tailed}} < 0.001$, $d = -2.50$. Furthermore, the adverb *rarely* did not merely modulate the plausibility of sentences containing typical and atypical objects but fully reversed the ratings, $t(23) = -5.136$, $p < 0.001$, $d = -1.05$.

ERPs—Overall the ERPs and patterns of effects in Experiment 3 were qualitatively similar to those observed in Experiment 2 (c.f. Figure 2 and Figure 3).

N400 (300-500ms): At the midline and mediolateral electrodes, the atypical object nouns elicited a large N400 in comparison with the typical nouns, largest over medial posterior electrodes (see Figure 3, Panels B and D; ANOVAs in Table 4). The main effect of quantifier type was not reliable and this factor did not interact with any factors of electrode location for midline or mediolateral electrodes. In the analysis of primary interest, the quantifier effects on the midline N400 amplitude for the typical and atypical object nouns were again in the expected directions (Figure 3, Panel C) and similar to those observed in Experiment 2 for the most- and few-type quantifiers. The largest (most negative) and smallest N400 amplitudes were observed in the context of *often* for the atypical and typical nouns, respectively. In the critical tests of the quantifier effect, in comparison with *often*, the adverb *rarely* increased the (negative going) N400 amplitude of the typical noun from 0.95 μV (*SD* = 1.71 μV) to 0.17 μV (*SD* = 1.36 μV), $t(23) = 1.78$, $p_{1\text{-tailed}} = 0.045$, $d = 0.36$. For the atypical noun, the N400 amplitude in the context of *rarely* (−1.538 μV, *SD* = 1.95 μV) was slightly lower than in the context of *often* (−1.81 μV, *SD* = 1.67 μV) though this numerical difference was not reliable ($p_{1\text{-tailed}} > 0.72$). Crucially, these small N400 amplitude modulations by the adverb *rarely* did not result in a crossover effect and the N400 for the atypical object noun remained reliably more negative than for the typical object noun, $t(23) = 3.77$, $p = 0.001$, $d = 0.77$. At the mediolateral electrodes, the pattern was generally similar to that observed at the midline electrodes. The N400 quantifier effects at the mediolateral electrodes were marginal for the typical objects ($p_{1\text{-tailed}} = 0.086$) and not reliable for the atypical objects ($p_{1\text{-tailed}} > 0.79$). Again, the N400 typicality effect did not crossover and atypical objects in the context of *rarely* (−1.08 μV, *SD* = 1.57) remained reliably more negative than for typical objects (0.289 μV, *SD* = 1.02), $t(23) = 3.94$, $p < 0.001$, $d = 0.81$.

LPC (500-800ms): At the midline electrodes, for the atypical nouns, potentials were more positive at the prefrontal electrode and more negative at the other locations and a similar pattern was observed at the mediolateral electrodes, where the atypical nouns were more positive at the prefrontal electrodes and more negative at medial posterior electrodes (ANOVAs in Table 4). No interaction effects involving the quantifier factor were reliable nor was the planned comparison at the left medial prefrontal electrode based on the maximum effect size observed in Experiment 2 ($p > 0.21$, $d = -0.26$).

SW: (800-1300ms): At midline and mediolateral electrodes, atypical nouns elicited an anterior positivity, largest at prefrontal electrodes (Table 4). This slow wave positivity was greatest at anterior electrodes for atypical nouns in the context of *rarely*, similar to the effect observed in Experiment 2 for the few-type quantifiers (c.f. Figures 2 and 3, Panel D). Based

on the effect size analysis in Experiment 2, planned comparisons were conducted for the three medial prefrontal electrodes and showed that in the context of *rarely*, atypical nouns were reliably more positive (~ 2 μV) than typical nouns. As in Experiment 2, the maximum effect size was observed at the left medial prefrontal electrode, 2.96 μV for atypical objects vs. 0.81 μV for typical, $t(23) = 4.11$, $p_{1\text{-tailed}} < 0.001$, $d = 0.84$, with large effects also observed at the midline prefrontal, 3.07 μV vs. 1.04 μV, $t(23) = 3.89$, $p_{1\text{-tailed}} < 0.001$, $d = 0.80$ and right medial prefrontal electrode, 3.36 μV vs. 1.02 μV, $t(23) = 3.94$, $p_{1\text{-tailed}} < 0.001$, $d = 0.80$. This prefrontal positivity for atypical vs. typical objects in the context of *rarely* was somewhat left lateralized with medium effect sizes ($d > 0.50$) at all left hemisphere electrodes except the two most posterior (occipital) channels. By comparison, over the right hemisphere, *d* exceeded 0.50 only at three medial frontal and prefrontal electrodes. There was no comparable object typicality effect in the context of the adverb *often*. The largest effect size observed for this comparison was small (midline occipital electrode, $d = 0.31$), smaller still at the three medial prefrontal electrodes, $0.24 \leq |d| \leq 0.29$ at the three medial prefrontal electrodes, and $d < 0.2$ at all other locations. This finding also patterns with Experiment 2, where the slow wave positivity effect associated with atypical nouns was restricted to the few-type quantifiers.

4.3 Experiment 3 Discussion

The pattern of plausibility judgments and ERPs under manipulations of the adverbs of quantification *often* and *rarely* and object noun typicality in Experiment 3 was qualitatively, and in most cases, quantitatively similar to the results obtained in Experiment 2 for the analogous manipulation of the most- and few-type determiners. The plausibility judgments in Experiment 3 demonstrated that the meaning of the expressions of quantity was appreciated and contributed to the global interpretation of the sentences in the expected way. In particular, the adverb *rarely* reverses the relative plausibility ratings for sentences containing typical and atypical object nouns though this crossover interaction is not completely symmetrical. Furthermore, the adverb *rarely* appears to exert less downward pressure on the plausibility of sentences containing typical object nouns than does the *few*-type determiner expressions employed in Experiment 2 which were given numerically lower plausibility ratings than in Experiment 3 and did not differ from the implausible sentences containing *most*-type quantifiers and atypical object nouns.

With respect to the ERPs, the effects of the quantifiers on typical and atypical object noun N400sand the prefrontal slow wave were of primary interest. Overall the general pattern of ERPs was very similar to Experiment 2. We conclude as before that quantifier semantics, in this case, adverbs of quantification, are registered incrementally and incorporated into the semantic context rapidly enough to have an impact on the processing of subsequent typical object nouns. In Experiment 2, the processing consequences of the subject determiner were evident on the object noun two words downstream, i.e., at a determiner-to-object-noun SOA of 1500ms. Experiment 3 sharpens this result by finding similar effects evident one word downstream at an adverb-to-object-noun SOA of 1000ms. In Experiment 3, we again find that the online and offline interpretive processes dissociate, replicating the pattern in Experiment 2: modulation of N400 amplitude by quantifiers in the expected direction but short of the crossover effect observed for plausibility judgments. This pattern suggests that the real-time interpretation of quantifier semantics is neither fully incremental nor entirely deferred, at least in these sorts of sentences (see general discussion). In addition, we again observed a prefrontal typicality effect following the N400 in the context of *rarely* quantifiers. This replication is a further indication of delayed and as yet not understood processing for *rarely* that is distinct from the processing associated with *often* and must be accounted for by any empirically adequate theory of real-time quantifier comprehension.

5. General Discussion

In a series of three RSVP reading experiments we tapped comprehender's background knowledge about agent-action-patient contingencies (Experiment 1) and then manipulated linguistic expressions of quantity to be consistent or inconsistent with this knowledge via quantified subject noun phrases, e.g., [*Most / Few*] *farmers grow* [*crops* / *worms*] (Experiment 2) and adverbs of quantification, e.g., *Farmers* [*often* / *rarely*] *grow* [*crops* / *worms*] (Experiment 3). In the latter two experiments we determined comprehender's interpretation of the quantifiers via post-sentence plausibility ratings and compared these offline judgments with the incremental interpretations inferred from on-line ERP measures of processing disruptions at the critical typical or atypical object noun. In Experiment 1, we found the predicted larger N400 amplitude for the atypical in comparison with typical object noun. In Experiment 2 we found small but reliable modulations of the typical object noun N400 amplitude as a function of most- vs. few-type quantified subject noun phrases and a similar pattern of N400 reductions was observed for the adverbs of quantification *often* vs. *rarely* in Experiment 3. Lexical factors that modulate N400 amplitude, e.g., length, frequency, and concreteness of the object noun are counterbalanced across quantifiers in this design, as are contextual factors such as lexical associations between the subject noun, main verb, and, object noun. We thus attribute modulation of the typical and atypical object noun N400 amplitudes to the contribution that the different quantifiers make to the evolving semantic context.

We take these N400 amplitude modulations as evidence of incremental quantifier interpretation and inconsistent with any hypothesis according to which the processing of quantifier semantics is entirely deferred or delayed. However, there is also an important dissociation between the patterns of quantifier and typicality effects for the offline and online measures. Whereas the quantifiers (*Most* vs. *Few* and *often* vs. *rarely)* reverse the offline plausibility judgments for sentences containing typical and atypical object nouns, they do *not* similarly reverse the N400 amplitudes for the object nouns. So although the ERP data indicate that quantifier meanings are registered in real-time and incrementally incorporated into the evolving representation of semantic context at least to some extent, these initial representations do not appear to be the same, more fully specified interpretations that inform the subsequent offline plausibility judgments. If this is correct, then at least in some respects, the semantic contributions of quantifier expressions to the interpretation of a sentence are processed at a delay and with a time course not yet fully understood.

We note that this interpretation depends essentially on the dissociation between the plausibility judgments and N400 amplitudes. These offline and online measures jointly afford an opportunity to draw sharper inferences than either the end-state sentence comprehension measures or the online ERP measures alone. Whereas the plausibility judgments provide evidence that the quantifiers are (eventually) fully interpreted, it would be a mistake to infer that they are fully interpreted at the time when the critical object noun is encountered. This is not to say that on-line measures are somehow more informative than off-line measures, for it would also be a mistake to conclude from on-line ERP evidence of underspecified quantifier interpretations that the quantifiers were not fully processed by sentence end (or ever). Rather, the conclusion that emerges—quantifiers are processed rapidly and incrementally though not fully when initially encountered, with full interpretations emerging later—is supported precisely by the dissociation between the off-line and on-line measures and cannot be drawn from either alone.

Our findings complement and, importantly, contrast with previous ERP investigations of quantifier interpretation. In a design that probed the resolution of referentially ambiguous

quantifier expressions Kaan, et al. (2007) manipulated the cardinality of bare quantifiers and found evidence of processing differences about a second later as a function of the number of objects already introduced into a simple discourse context. Our design does not essentially involve ambiguity resolution or intra-sentential discourse reference but rather examines the contribution of quantifier information to the sentential semantic context that evolves within isolated sentences. In this respect our design has more in common with Kounios & Holcomb (1992). There are a number of differences between their study and ours and perhaps the most salient concerns the results: we observed N400 evidence that the quantified subject noun phrases modulate processing of the object noun whereas Kounios & Holcomb (1992) did not. Our findings thus appear to be inconsistent with the suggestion that N400 primarily reflects aspects of the organization of semantic memory to the exclusion of structural semantic factors.

An unexpected additional finding in these experiments is a prefrontal slow wave positivity for atypical vs. typical object nouns. We found this object noun typicality effect to be most pronounced in the context of the *few*-type quantifiers (Experiment 2) and adverb *rarely* (Experiment 3). The time course suggests that these constructions require additional or secondary processing, perhaps related to interpretation (resolving explicit or implicit negatives?) or related to the comparison with background knowledge or decision processes relevant to the plausibility judgment. Positivities evolving after the N400 have been widely observed in ERP sentence comprehension research. Variously termed P600 (Osterhout & Holcomb, 1992) and Syntactic Positive Shift (Hagoort, Brown, & Groothusen, 1993), these effects are often largest over posterior scalp and associated with grammatical disruptions, e.g., words that violate grammatical rules or that are inconsistent with the preferred interpretation of a structural ambiguity. The relation between the frontal positivities observed in our experiments where there is no obvious syntactic ambiguity or anomaly and the many previously reported late posterior positivities is unclear and the relation between semantic and syntactic processing and the negative and positive waveforms that emerge between about 300ms and 1200ms poststimulus is not simple (for reviews and critical discussion see Bornkessel-Schlesewsky & Schlesewsky, 2008; Kolk & Chwilla, 2007; Kuperberg, 2007).

There are, however, a few reports of late positivities with a predominantly frontal distribution in experimental designs that, like ours, do not involve grammatical or structural disruptions. Moreno, Federmeier, & Kutas, 2002 found that for Spanish-English bilinguals reading English sentences and idioms, a late frontal positivity (650 – 850 ms) was elicited both by unexpected English completions (lexical switches) as well as Spanish translations of the expected English completion (code switches), particularly for the idioms. There is also some preliminary evidence of differential involvement of the cerebral hemispheres. In a study of metaphor comprehension that included literal controls Coulson & Van Petten, 2007 also observed a late anterior positivity (600-900ms) for plausible but unexpected (low cloze) sentence final words in comparison with the expected (high cloze) endings, though only for words presented in the right-hemifield (left hemisphere). Further evidence and, importantly, a clear dissociation between the late positivity and the N400 is reported by Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007. Words may be more or less expected in context with expectancy operationalized via cloze probability, i.e., the probability of production in an offline sentence completion task. Sentence contexts may be more or less constraining where constraint is defined as the highest cloze value of the completions. Replicating Kutas & Hillyard (1984) they found that low cloze sentence final words elicited a larger N400 than high cloze and, furthermore, that for low cloze words, there was no effect of sentential constraint on the N400 amplitude. That is, unexpected words had similar N400s regardless of whether they were unexpected alternatives to a highly expected word or unexpected because the sentential context did not provide enough information to generate strong

expectations. However, there was pronounced frontal slow wave positivity when these unexpected words occurred in highly constraining contexts in comparison with weakly constraining contexts. The authors suggest this prefrontal positivity may reflect the appreciation of a mismatch between the expected item and the word presented or the allocation of resources necessary to override or revise a prediction or both.

It is difficult to see how this line of reasoning can be extended to the pattern of data in our Experiment 2 and Experiment 3. Whatever the space of expected continuations might be for the most-type quantifier or *often* sentence contexts, e.g., *most farmers grow ___*, in the absence of a supporting discourse context, it is difficult to generate strong expectancies about the continuation of *Few farmers grow ___*. In our experiment, if anything, those sentences containing the few-type quantifiers and adverb *rarely* should be less constraining than those sentences with the most-type quantifiers and *often*. If a prefrontal slow wave positivity is associated with unexpected words in high vs. low constraint contexts we would expect to see the clearest evidence at the atypical object noun *worms*, in *Most farmers grow worms* in comparison with *Few farmers grow worms* (or perhaps in comparison with *Few farmers grow crops*, the question of which control is appropriate is debatable, though less critical since either choice should be relatively less positive by comparison). Although the prefrontal positivity was indeed greatest for the word *worms*, it occurred in the less constraining sentential contexts that contained the few-type quantifiers and the adverb *rarely*. There are many possible explanations for these discrepant findings. It may be that qualitatively similar prefrontal positivities reflect different functional processes in the two experiments. Alternatively, the prefrontal positivity may reflect a process that is common to both, e.g., allocation of processing resources as proposed by Federmeier, et al., 2007 though contra their suggestion, not specifically linked to the revision of a prediction. In addition, plausibility may be playing a different role in the two cases. In our experiment the atypical noun, *worms*, in the context of the few-type quantifiers although unexpected is, based on the response data, ultimately plausible. Further investigation is needed to determine whether the frontal positivity reflects processing selectively associated with the few- in contrast with most-type quantifiers or some aspect of the plausibility evaluation triggered in this experiment.

Finally, in evaluating the generalizability of our quantifier results we are alive to a potentially instructive parallel with recent research on the real-time comprehension of negation. It is uncontroversial that negation contributes to the overall semantics of a sentence. Although the Fischler, et al. (1983) report that negation did not have a reliable effect on N400 amplitude of the predicate term in simple subject-predicate sentences appears to militate against incremental interpretation of negation (see also Kounios & Holcomb, 1992; Ludtke, Friedrich, De Filippis, & Kaup, 2008), the scope of this result has been sharply circumscribed by recent evidence that negation can be processed incrementally when it is pragmatically licensed by the context. In isolated sentences, explicit denials may provide little useful information, e.g., *A robin is not a tree*, although true, is uninformative and thus pragmatically infelicitous. However, against the backdrop of appropriate contexts, denials may be highly informative, for example, when a speaker attempts to correct a listener's mistaken belief as in, *A robin is not a member of the finch family*. In recent work, Staab (2007) and Nieuwland & Kuperberg (2008) independently found that N400 amplitude on critical target words varied in a manner consistent with the incremental interpretation of negation, provided it was pragmatically supported (licensed) by contextual information (c.f., Wason, 1965). Nieuwland & Kuperberg (2008) recorded ERPs in sentences such as, *With proper equipment, scuba diving is very* [*safe / dangerous*], and found that N400 amplitude for *dangerous* was greater than for *safe* and, crucially, also found that this relationship reversed when the copula *is* was replaced by *isn't*. This result, in conjunction with their other comparisons was taken as evidence for the incremental interpretation of negation. In our

quantifier experiments, we observed N400 amplitude modulation but not reversal at critical target words as a function of quantifier type and are suggesting that this is evidence of incremental construction of partially specified quantifier interpretations. As noted above, our few-type quantifiers are "negative" in the sense that they license negative polarity items. Whether or not the semantics of these quantifier expressions, by analogy with explicit negation markers, might be interpreted incrementally and fully in pragmatically supporting contexts is an open question.

6. Conclusion

We investigated the real-time processing of non-logical quantifier expressions such as *most, few, often*, and *rarely*, in sentences that activated background knowledge about typical and atypical agent-action-patient contingencies. We found evidence from off-line plausibility ratings that the quantifier expressions were interpreted fully, i.e., consistent with theoretical accounts of their meaning. We also found on-line evidence from N400 amplitude modulations at critical target words that the meanings of quantifiers were initially registered but according to the offline plausibility judgments were not fully incorporated into the evolving representation of the sentential semantic context. Our findings are thus inconsistent with two types of models: those on which quantifier interpretation is immediate and full and those on which quantifier interpretation is entirely delayed until lexico-semantic and subject-predicate relationships are established. We offer these results as evidence for the incremental partial interpretation of these quantifier expressions, with the full interpretation (inferred from the plausibility judgments) occurring at some delay. A number of open questions remain to be investigated. Different patterns of N400 effects and frontal slow wave positivities were observed at the typical and atypical object nouns, predominantly in the context of the few-type and *rarely* quantifier expressions. This result raises the possibility that there may be systematic processing differences for the two types of quantifiers and if so, no undifferentiated model of quantifier interpretation will be empirically adequate. Finally, even if incremental interpretations of quantifiers are underspecified in isolated sentences, it remains to be determined whether pragmatically supporting discourse contexts might result in not only immediate but full quantifier interpretation.

Acknowledgments

This research was supported by grant NICHD22614 and AG08313 to M.K. We thank Julia Cagle and Jessica Bradshaw for their assistance with data collection and are grateful to Kara Federmeier, Katherine DeLong, and three anonymous reviewers for helpful discussions and comments on previous versions of this report.

Appendix: Stimuli

Sentences in the two conditions in Experiment 1 began with the bare plural subject noun and contained the typical or atypical object nouns (underlined). The four conditions in Experiment 2 crossed the most-type and few-type quantifiers (italics) with the typical and atypical object nouns. The four conditions in Experiment 3 began with the bare plural subject noun followed by *often* or *rarely* crossed with the typical and atypical object nouns.

1. *Most/Few* prosecutors accuse defendants/sheriffs of committing a crime.
2. *Almost all/Almost no* groupies follow singers/boys around the country.
3. *Practically all/Practically no* postmen carry mail/oil in their satchel.
4. *A large number of/A small number of* kittens chase mice/flies if given the chance.
5. *Nearly all/Rather few* doctors treat cancer/hunger as a serious condition.

6. *Lots of/Hardly any* chefs cook pancakes/pears for breakfast at roadside diners.
7. *A lot of/A very few* boxers fight opponents/lawsuits in cities where there are large crowds.
8. *Many/Few* archaeologists find artifacts/gold when they excavate ancient cities.
9. *Most/Few* farmers grow crops/worms as their primary source of cash.
10. *Almost all/Almost no* veterinarians help animals/coyotes when they get injured.
11. *Practically all/Practically no* lions hunt gazelles/bears as they prowl the savanah.
12. *A large number of/A small number of* brides kiss grooms/babies during their wedding ceremony.
13. *Nearly all/Rather few* artists paint portraits/maps to pay their bills.
14. *Lots of/Hardly any* historians study articles/poetry written by their colleagues.
15. *A lot of/A very few* authors write books/lists to make a living.
16. *Many/Few* bands record albums/hits throughout their careers.
17. *Most/Few* cats scratch furniture/walls if they are not declawed.
18. *Almost all/Almost no* janitors clean floors/dishes after cleaning everything else.
19. *Practically all/Practically no* lawyers argue cases/politics in a court of law.
20. *A large number of/A small number of* judges sentence murderers/monks to life in prison.
21. *Nearly all/Rather few* waitresses bring cocktails/napkins to customers who get their attention.
22. *Lots of/Hardly any* hunters shoot game/livestock during the hunting season.
23. *A lot of/A very few* theaters show movies/operas on Saturday nights.
24. *Many/Few* butchers slaughter cows/ducks and sell the meat.
25. *Most/Few* pickpockets steal wallets/shoes while the owner is distracted.
26. *Almost all/Almost no* psychics read fortunes/news for their customers every day.
27. *Practically all/Practically no* pilots fly planes/kites on cloudy days.
28. *A large number of/A small number of* thieves take jewels/jobs from their victims when they can.
29. *Nearly all/Rather few* professors teach classes/friends while doing research.
30. *Lots of/Hardly any* pitchers throw balls/rocks to warm up for a game.
31. *A lot of/A very few* kidnappers blindfold captives/infants to keep their identities secret.
32. *Many/Few* smugglers transport drugs/umbrellas into the United States.
33. *Most/Few* tourists visit museums/mines on their vacation.
34. *Almost all/Almost no* sentries guard bases/dumpsters at all hours of the night.
35. *Practically all/Practically no* telemarketers call people/inmates to tell them about new products.

36. *A large number of/A small number of* punters kick footballs/boxes for practice on their days off.

37. *Nearly all/Rather few* shoppers buy groceries/ammunition at the local supermarket.

38. *Lots of/Hardly any* speakers address audiences/letters using a microphone.

39. *A lot of/A very few* armies battle invaders/immigrants and try to drive them out of the country.

40. *Many/Few* gardeners plant flowers/bombs to make their garden prettier.

41. *Most/Few* architects design buildings/closets with lots of windows.

42. *Almost all/Almost no* negotiators settle disputes/payments as quickly as possible.

43. *Practically all/Practically no* witnesses describe robbers/tellers as being large and menacing.

44. *A large number of/A small number of* couriers deliver packages/organs sealed in boxes.

45. *Nearly all/Rather few* brokers sell stocks/curtains in after-hours trading.

46. *Lots of/Hardly any* mechanics fix cars/toys that have broken parts.

47. *A lot of/A very few* joggers run laps/monthly to stay in shape.

48. *Many/Few* cleaners wash clothes/rags using strong laundry soap.

49. *Most/Few* squirrels gather nuts/nails and store them for the winter.

50. *Almost all/Almost no* investors loan money/papers to businesses which are profitable.

51. *Practically all/Practically no* morticians arrange funerals/parties after a lot of planning.

52. *A large number of/A small number of* actors perform plays/surgery on stage.

53. *Nearly all/Rather few* adolescents play games/horns all through high school.

54. *Lots of/Hardly any* cowboys ride horses/bicycles when they go to the back country.

55. *A lot of/A very few* barbers cut hair/steak while carrying on a conversation.

56. *Many/Few* bartenders mix drinks/metaphors while talking to the patrons.

57. *Most/Few* executives make decisions/models that reflect their business acumen.

58. *Almost all/Almost no* warehouses store merchandise/sugar which has not been sold yet.

59. *Practically all/Practically no* coaches train athletes/soldiers to work at a high level of performance.

60. *A large number of/A small number of* plumbers remove clogs/weeds using special equipment.

61. *Nearly all/Rather few* satirists poke fun/knives at famous people.

62. *Lots of/Hardly any* nurses draw blood/sketches for medical testing and analysis.

63. *A lot of/A very few* caterers prepare food/baths for large groups of people.

64. *Many/Few* weathermen report storms/weddings that are expected in the next few days.

65. *Most/Few* knights rescue damsels/dragons from the clutches of an ogre.

66. *Almost all/Almost no* generals command troops/civilians in time of war.

67. *Practically all/Practically no* boxcars hold cargo/feathers that can take extreme temperatures.

68. *A large number of/A small number of* magazines publish stories/songs that involve current events.

69. *Nearly all/Rather few* policemen arrest crooks/priests after the bars close down.

70. *Lots of/Hardly any* actresses wear dresses/helmets that were created by famous designers.

71. *A lot of/A very few* adults eat chicken/grass on a regular basis.

72. *Many/Few* teachers punish students/aides that are disrupting the class.

73. *Most/Few* governments build monuments/trains during times of national importance.

74. *Almost all/Almost no* scientists conduct research/traffic with great attention to detail.

75. *Practically all/Practically no* comedians entertain crowds/politicians at comedy clubs around the country.

76. *A large number of/A small number of* instructors evaluate pupils/grapes by giving them a test.

77. *Nearly all/Rather few* supervisors discipline workers/shareholders who show up late.

78. *Lots of/Hardly any* patrolmen question suspects/minors to get information about a crime.

79. *A lot of/A very few* parents lecture children/pets about not playing too roughly.

80. *Many/Few* bellboys lug suitcases/fruits every day at work.

81. *Most/Few* businessmen employ accountants/ministers to keep track of the books.

82. *Almost all/Almost no* kids want candy/peas for dessert after dinner.

83. *Practically all/Practically no* runners drink water/tea while competing in a marathon.

84. *A large number of/A small number of* ranches hire hands/drunks to take care of the chores.

85. *Nearly all/Rather few* countries erect statues/bridges to honor their national heroes.

86. *Lots of/Hardly any* clowns toss pies/cookies at each other as part of their act.

87. *A lot of/A very few* ranchers feed cattle/visitors out on the open prairie.

88. *Many/Few* automobiles need gas/sunlight to run for any length of time.

89. *Most/Few* newlyweds receive gifts/bills at their wedding reception.

90. *Almost all/Almost no* sages offer advice/sweets to people who ask them for their opinion.

91. *Practically all/Practically no* attorneys meet clients/dates at their law firm.

92. *A large number of/A small number of* engineers plan projects/meals for wealthy executives.

93. *Nearly all/Rather few* bakers slice bread/pizza in a special cutting machine.

94. *Lots of/Hardly any* retirees trust banks/strangers with their life's savings.

95. *A lot of/A very few* quarterbacks lift weights/wheels as part of their training regimen.

96. *Many/Few* psychologists use hypnosis/violence to help patients remember their childhood.

97. *Most/Few* matadors wave capes/flags to goad the bull into attacking.

98. *Almost all/Almost no* eskimos catch fish/malaria during the long arctic summer.

99. *Practically all/Practically no* astronomers observe stars/insects that have unusual properties.

100. *A large number of/A small number of* pirates ransom prisoners/husbands for gold or jewels.

101. *Nearly all/Rather few* employees do work/crosswords while they are at the office.

102. *Lots of/Hardly any* hosts invite guests/butlers to stay for dinner.

103. *A lot of/A very few* surgeons request specialists/bodyguards for particularly difficult operations.

104. *Many/Few* chimps peel bananas/apples before they eat them.

105. *Most/Few* mayors see citizens/ghosts on a regular basis.

106. *Almost all/Almost no* dogs gnaw bones/tires to exercise their jaws.

107. *Practically all/Practically no* sailors abandon ships/rowboats that are about to capsize.

108. *A large number of/A small number of* spies collect data/sand from foreign countries.

109. *Nearly all/Rather few* songwriters create music/stages specifically for their own band.

110. *Lots of/Hardly any* snakes devour eggs/pigs in one large bite.

111. *A lot of/A very few* detectives notice clues/reporters all around the crime scene.

112. *Many/Few* families schedule vacations/discussions during the summer months.

113. *Most/Few* therapists make diagnoses/deals based on years of experience.

114. *Almost all/Almost no* lifeguards protect swimmers/birds at the beach.

115. *Practically all/Practically no* dentists pull teeth/files in their office.

116. *A large number of/A small number of* salesmen market products/values to prospective buyers.

117. *Nearly all/Rather few* aquariums give sharks/donors a lot of special attention.

118. *Lots of/Hardly any* burglars have disguises/hostages when they are breaking into a house.

119. *A lot of/A very few* tenants rent apartments/lawnmowers on a monthly basis.

120.*Many/Few* custodians scrub sinks/trucks before applying a disinfectant.

References

Altmann GTM, Mirkovic J. Incrementality and prediction in human sentence processing. Cognitive Science 2009;33:583–609. [PubMed: 20396405]

Bah, T. Inkscape: guide to a vector drawing program. Prentice Hall Press; 2007.

Barton SB, Sanford AJ. A case-study of anomaly detection: Shallow semantic processing and cohesion establishment. Memory & Cognition 1993;21:477–487.

Barwise J, Cooper R. Generalized quantifiers and natural-language. Linguistics and Philosophy 1981;4:159–219.

Bornkessel-Schlesewsky I, Schlesewsky M. An alternative perspective on “semantic P600” effects in language comprehension. Brain Research Reviews 2008;59:55–73. [PubMed: 18617270]

Carpenter PA, Just MA. Sentence comprehension: Psycholinguistic processing model of verification. Psychological Review 1975;82:45–73.

Christianson K, Hollingworth A, Halliwell JF, Ferreira F. Thematic roles assigned along the garden path linger. 2001

Clark HH, Chase WG. Process of comparing sentences against pictures. Cognitive Psychology 1972;3:472–517.

Cohen, J. Statistical power analysis for the behavioral sciences. 2nd ed.. Lawrence Earlbaum Associates; Hillsdale, NJ: 1988.

Coulson S, Van Petten C. A special role for the right hemisphere in metaphor comprehension? ERP evidence from hemifield presentation. 2007

Dale, A. Source Localization and Spatial Discriminant Analysis of Event-Related Potentials: Linear Approaches. Vol. 55. University of California; San Diego: 1994. p. 2559Doctoral dissertation1994). Dissertation Abstracts International

Diesing M. Bare plural subjects and the derivation of logical representations. Linguistic Inquiry 1992;23:353–380.

Erickson TD, Mattson ME. From words to meaning: A semantic illusion. Journal of Verbal Learning and Verbal Behavior 1981;20:540–551.

Fauconnier G. Pragmatic Scales and Logical Structure. Linguistic Inquiry 1975;6:353–375.

Federmeier KD, Wlotko EW, De Ochoa-Dewald E, Kutas M. Multiple effects of sentential constraint on word processing. Brain Research 2007;1146:75–84. [PubMed: 16901469]

Ferguson HJ, Sanford AJ. Anomalies in real and counterfactual worlds: An eye-movement investigation. Journal of Memory and Language 2008;58:609–626.

Ferguson HJ, Sanford AJ, Leuthold H. Eye-movements and ERPs reveal the time course of processing negation and remitting counterfactual worlds. Brain Research 2008;1236:113–125. [PubMed: 18722356]

Ferreira F, Bailey KGD, Ferraro V. Good-enough representations in language comprehension. Current Directions in Psychological Science 2002;11:11–15.

Filik R. Contextual override of pragmatic anomalies: Evidence from eye movements. Cognition 2008;106:1038–1046. [PubMed: 17524387]

Filik R, Paterson KB, Liversedge SP. Processing doubly quantified sentences: Evidence from eye movements. Psychonomic Bulletin & Review 2004;11:953–959. [PubMed: 15732709]

Fischler I, Bloom PA, Childers DG, Roucos SE, Perry NW. Brain potentials related to stages of sentence verification. Psychophysiology 1983;20:400–409. [PubMed: 6356204]

Frazier L, Clifton C, Rayner K, Deevy P, Koh S, Bader M. Interface problems: Structural constraints on interpretation? Journal of Psycholinguistic Research 2005;34:201–231. [PubMed: 16050443]

Frazier L, Rayner K. Taking on semantic commitments: Processing multiple meanings vs. multiple senses. Journal of Memory and Language 1990;29:181–200.

Ganis G, Kutas M, Sereno MI. The search for “common sense”: An electrophysiological study of the comprehension of words and pictures in reading. Journal of Cognitive Neuroscience 1996;8:89–106.

Greenhouse SW, Geisser S. ON METHODS IN THE ANALYSIS OF PROFILE DATA. Psychometrika 1959;24:95–112.

Hagoort P, Brown C, Groothusen J. The syntactic positive shift (SPS) as an ERP measure of syntactic processing. Language and Cognitive Processes 1993;8:439–483.

Hagoort P, Hald L, Bastiaansen M, Petersson KM. Integration of word meaning and world knowledge in language comprehension. Science 2004;304:438–441. [PubMed: 15031438]

Hagoort P, van Berkum J. Beyond the sentence given. Philosophical Transactions of the Royal Society B-Biological Sciences 2007;362:801–811.

Herron, T. cleave (Version January 30, 2005). Human Cognitive Neurophysiology Laboratory, University of California; Davis: 2005.

Just MA, Carpenter PA. A theory of reading: From eye fixations to comprehension. Psychological Review 1980;87:329–354. [PubMed: 7413885]

Kaan E, Dallas AC, Barkley CM. Processing bare quantifiers in discourse. 2007

Keenan EL, Stavi J. A semantic characterization of natural-language determiners. Linguistics and Philosophy 1986;9:253–326.

Kolk H, Chwilla D. Late positivities in unusual situations. Brain and Language 2007;100:257–261. [PubMed: 16919324]

Kounios J, Holcomb PJ. Structure and process in semantic memory: Evidence from event-related brain potentials and reaction-times. Journal of Experimental Psychology-General 1992;121:459–479. [PubMed: 1431739]

Krifka M. The semantics and pragmatics of polarity items. Linguistic Analysis 1995;25:209–258.

Kucera, H.; Francis, WN. Computational analysis of present-day American English. Brown University Press; Providence: 1967.

Kuperberg GR. Neural mechanisms of language comprehension: Challenges to syntax. Brain Research 2007;1146:23–49. [PubMed: 17400197]

Kurtzman HS, Macdonald MC. Resolution of quantifier scope ambiguities. Cognition 1993;48:243–279. [PubMed: 8269698]

Kutas M, Hillyard SA. Reading sensless sentences: Brain potentials reflect semantic incongruity. Science 1980;207:203–205. [PubMed: 7350657]

Kutas M, Hillyard SA. Brain potentials during reading reflect word expectancy and semantic association. Nature 1984;307:161–163. [PubMed: 6690995]

Kutas, M.; Van Petten, CK.; Kluender, R. Psycholinguistics Electrified II (1994-2005). In: Traxler, MJ.; Gernsbacher, MA., editors. Handbook of Psycholinguistics. 2nd ed.. Academic Press; Amsterdam: 2006. p. 659-724.

Lewis, D. Adverbs of quantification. In: Keenan, EL., editor. Formal semantics of natural language. Cambridge University Press; Cambridge: 1975. p. 3-15.

Ludtke J, Friedrich CK, De Filippis M, Kaup B. Event-related potential correlates of negation in a sentence-picture verification paradigm. Journal of Cognitive Neuroscience 2008;20:1355–1370. [PubMed: 18303972]

Moreno EM, Federmeier KD, Kutas M. Switching languages, switching palabras (words): An electrophysiological study of code switching. Brain and Language 2002;80:188–207. [PubMed: 11827443]

Moxey LM, Filik R, Paterson KB. On-line effects of what is expected on the resolution of plural pronouns. Language and Cognitive Processes 2009;24:843–875.

Moxey LM, Sanford AJ, Dawydiak EJ. Denials as controllers of negative quantifier focus. Journal of Memory and Language 2001;44:427–442.

Nieuwland MS, Kuperberg GR. When the truth is not too hard to handle: An event-related potential study on the pragmatics of negation. Psychological Science 2008;19:1213–1218. [PubMed: 19121125]

Nieuwland MS, Van Berkum JJA. When peanuts fall in love: N400 evidence for the power of discourse. Journal of Cognitive Neuroscience 2006;18:1098–1111. [PubMed: 16839284]

Osterhout L, Holcomb PJ. Event-related brain potentials elicited by syntactic anomaly. Journal of Memory and Language 1992;31:785–806.

Paterson KB, Filik R, Liversedge SP. Competition during the processing of quantifier scope ambiguities: Evidence from eye movements during reading. Quarterly Journal of Experimental Psychology 2008;61:459–473.

R Development Core Team. R: A Language and Environment for Statistical Computing (Version 2.9.0 (2009-04-17)). R Foundation for Statistical Computing; Vienna, Austria: 2009.

Rayner K. Eye-movements in reading and information-processing. Psychological Bulletin 1978;85:618–660. [PubMed: 353867]

Rayner K, Clifton C. Language processing in reading and speech perception is fast and incremental: Implications for event-related potential research. Biological Psychology 2009;80:4–9. [PubMed: 18565638]

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.

Reder LM, Kusbit GW. Locus of the Moses illusion: Imperfect encoding, retrieval, or match. Journal of Memory and Language 1991;30:385–406.

Sanford AJ, Dawydiak EJ, Moxey LM. A unified account of quantifer perspective effects in discourse. Discourse Processes 2007;44:1–32.

Sanford AJ, Sturt P. Depth of processing in language comprehension: Not noticing the evidence. Trends in Cognitive Sciences 2002;6:382–386. [PubMed: 12200180]

Staab, J. Negation in context: Electrophysiological and behavioral investigations of negation effects in discourse processing. La Jolla; University of California, San Diego and San Diego State: 2007.

Sturt P, Sanford AJ, Stewart A, Dawydiak E. Linguistic focus and good-enough representations: An application of the change-detection paradigm. Psychonomic Bulletin & Review 2004;11:882–888. [PubMed: 15732698]

Trabasso T, Rollins H, Shaughnessy E. Storage and verification stages in processing concepts. Cognitive Psychology 1971;2:239–289.

Tukey, JW. Exploratory Data Analysis. Addison-Wesley Publishing Company; Reading, MA: 1977.

Ward P, Sturt P. Linguistic focus and memory: An eye movement study. Memory & Cognition 2007;35:73–86.

Warren T, McConnell K. Investigating effects of selectional restriction violations and plausibility violation severity on eye-movements in reading. Psychonomic Bulletin & Review 2007;14:770–775. [PubMed: 17972747]

Warren T, McConnell K, Rayner K. Effects of context on eye movements when reading about possible and impossible events. Journal of Experimental Psychology-Learning Memory and Cognition 2008;34:1001–1010.

Wason PC. The contexts of plausible denial. Journal of Verbal Learning and Verbal Behavior 1965;4:7–11.

Wickham, H. ggplot2: Elegant graphics for data analysis. Springer; New York: 2009.

Wijnen F, Kaan E. Dynamics of semantic processing: The interpretation of bare quantifiers. Language and Cognitive Processes 2006;21:684–720.

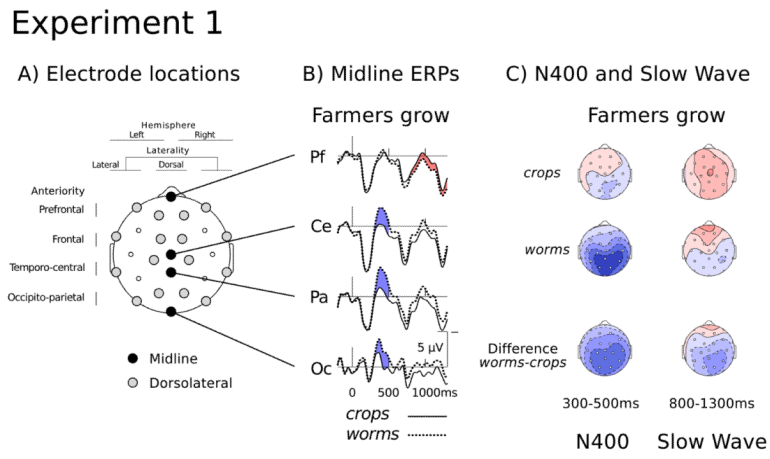

Figure 1.

Experiment 1. A) Electrode locations and ANOVA factors. Black circles indicate the four midline electrode locations, gray circles indicate the 16 mediolateral electrodes analyzed. B) Midline ERPs timelocked to the onset of the critical object nouns. In these and subsequent figures, negative is plotted up, waveforms are low-pass filtered at 10Hz for graphical representation, the N400 effect (300-500ms) is shaded in blue, the prefrontal Slow Wave effect (800-1300ms) is shaded in red. C) Spline interpolated maps of the scalp potential distributions for the object noun N400s, Slow Waves, and effects (differences). In these and subsequent figures, each isopotential contour spans 0.625μV with more negative potentials darker shades of blue and more positive potentials darker shades of red.

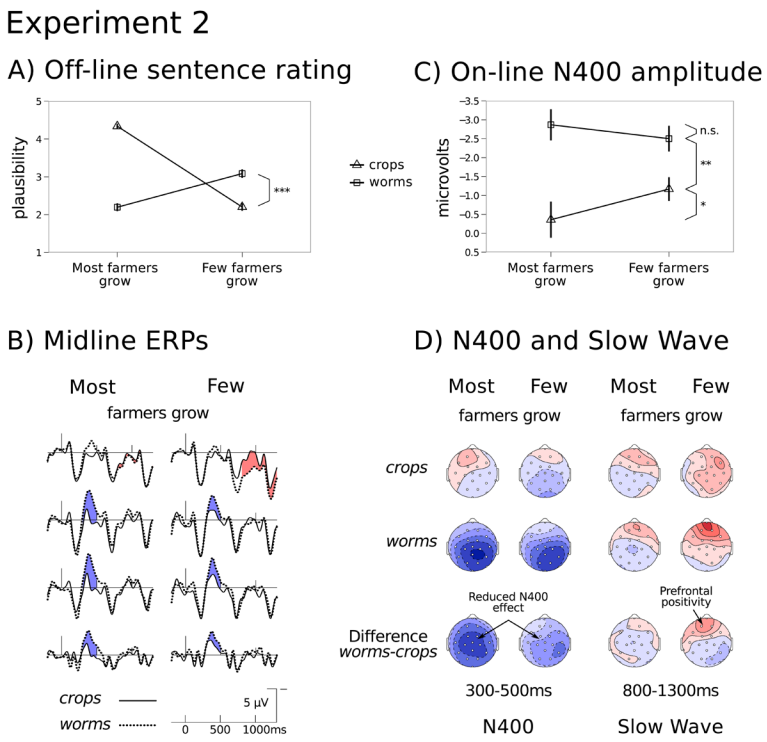

Figure 2.

Experiment 2. A) Mean plausibility ratings (1=highly implausible, 2=moderately implausible, 3=neutral, 4=moderately plausible, 5=highly plausible) recorded following the presentation of each sentence (vertical bars = 1 *SE*). B) Midline ERPs timelocked to the onset of the critical object nouns. C) Mean N400 amplitude pooled across central and posterior midline electrodes for the critical object nouns (vertical bars = 1 *SE*). D) Spline interpolated maps of the scalp potential distributions for the object noun N400s, Slow Waves, and effects (differences). In these and subsequent figures effects marked with * indicates $p_{1\text{-tailed}} < 0.05$. Effects marked with ** and *** indicate $p_{2\text{-tailed}} < 0.01$ and 0.001 respectively. Effects marked n.s. are not statistically significant.

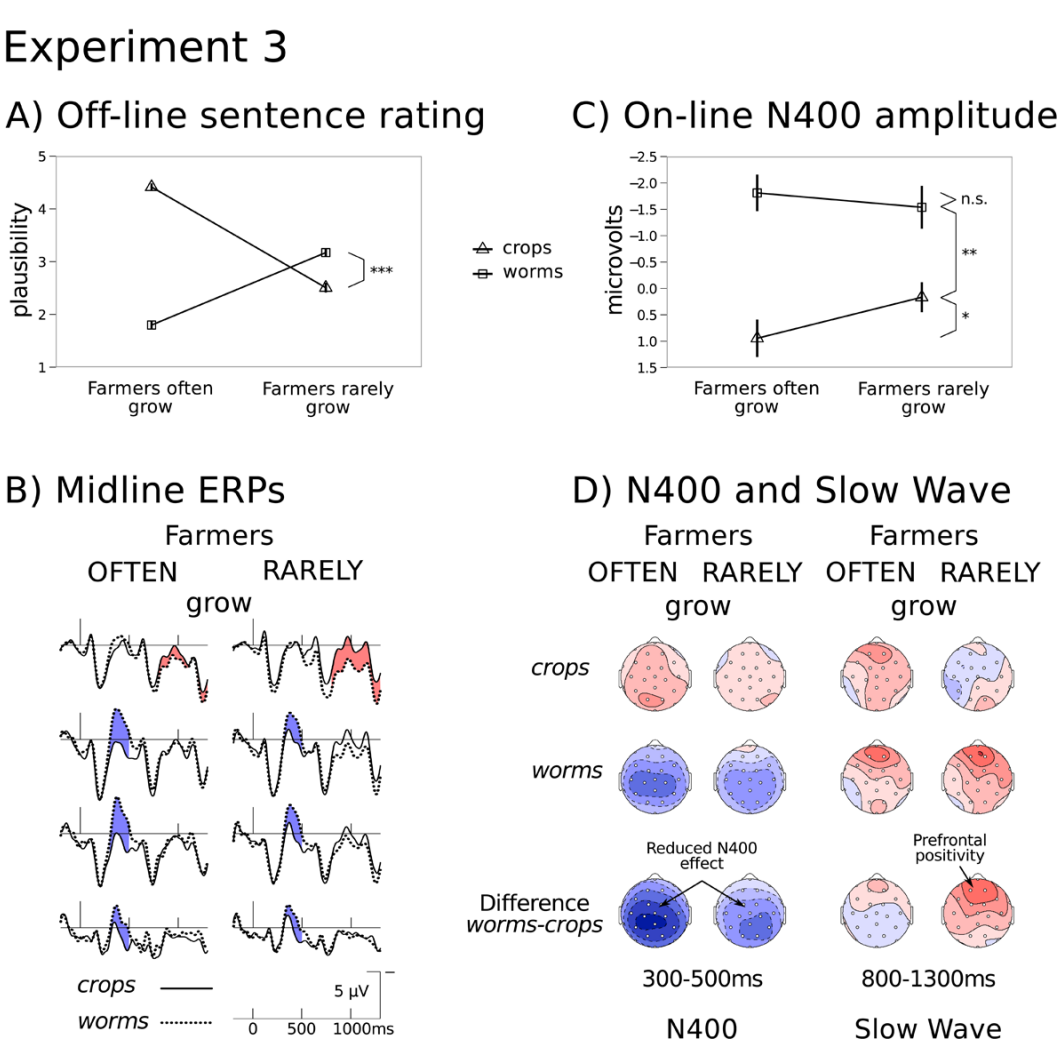

Figure 3.

Experiment 3. A) Mean plausibility ratings recorded following the presentation of each sentence (vertical bars = 1 *SE*). B) Midline ERPs timelocked to the onset of the critical object nouns. C) Mean N400 amplitudes pooled across central and posterior midline electrodes for the critical object nouns (vertical bars = 1 *SE*). D) Spline interpolated maps of the scalp potential distributions for the object noun N400s, Slow Waves, and effects (differences).

Table 1

Example sentences

	Condition		Example sentence
	Quantifier	Object	
Experiment 1	bare plural	typical	*Farmers* grow crops as their primary source of income
	bare plural	atypical	*Farmers* grow worms as their primary source of income
Experiment 2	most-type	typical	*Most* farmers grow crops as their primary source of income
	most-type	atypical	*Most* farmers grow worms as their primary source of income
	few-type	typical	*Few* farmers grow crops as their primary source of income
	few-type	atypical	*Few* farmers grow worms as their primary source of income
Experiment 3	often	typical	Farmers *often* grow crops as their primary source of income
	often	atypical	Farmers *often* grow worms as their primary source of income
	rarely	typical	Farmers *rarely* grow crops as their primary source of income
	rarely	atypical	Farmers *rarely* grow worms as their primary source of income

Note. Quantifier expressions are in italics and object nouns are underlined here for expository purposes and were not so marked in the experiments.

Table 2

Experiment 1 ERP Analysis of Variance

Source	*df*	*F*	*p*	*MSE*	η_p^2	ε_{GG}
	Midline electrodes					
300-500ms						
T	1, 31	26.79	< 0.001	5.92	0.46	n.a.
T × A	3, 93	13.87	< 0.001	0.95	0.31	0.603
500-800ms						
T × A	3, 93	8.22	0.002	1.12	0.21	0.561
800-1300ms						
T × A	3, 93	8.97	0.002	1.20	0.22	0.480
	Mediolateral electrodes					
300-500ms						
T	1, 31	26.89	< 0.001	16.95	0.46	n.a.
T × L	1, 31	21.15	< 0.001	2.49	0.41	n.a.
T × A	3, 93	7.58	0.006	2.01	0.20	0.399
T × H × L × A	3, 93	3.68	0.026	0.11	0.11	0.748
500-800ms						
T × L	1, 31	5.36	0.027	3.39	0.15	n.a.
T × A	3, 93	5.02	0.025	2.26	0.14	0.40
T × H × A	3, 93	5.22	0.011	0.31	0.14	0.582
T × H × L × A	3, 93	4.18	0.012	0.13	0.12	0.768
800-1300ms						
T × L	1, 31	5.71	0.023	2.45	0.16	n.a.
T × A	3, 93	7.57	0.025	2.60	0.14	0.413
T × L × A	3, 93	5.35	< 0.001	0.55	0.15	0.638

Note: In this and subsequent ANOVA tables, the factor abbreviations are as follows: T = typicality, H = Hemisphere, L = Laterality, A = Anteriority; degrees of freedom are for the numerator, denominator; only effects reliable at alpha = .05 are listed.

Table 3

Experiment 2 ERP Analysis of Variance

Source	*df*	*F*	*p*	*MSE*	η_p^2	ε_{GG}
	Midline electrodes					
300-500ms						
T	1, 19	50.17	< 0.001	5.16	0.73	n.a.
	Mediolateral electrodes					
50-150ms						
T × H × A	3, 57	3.63	0.041	0.21	0.16	0.602
T × H × L × A	3, 57	4.65	0.011	0.09	0.20	0.771
300-500ms						
T	1, 19	45.47	< 0.001	15.23	0.71	n.a.
T × L	1, 19	12.27	0.002	4.03	0.39	n.a.
T × H × A	3, 57	3.56	0.043	0.25	0.16	0.613
Q × T × L	1, 19	5.45	0.031	1.63	0.22	n.a.
T × H × L × A	3, 57	8.10	0.001	0.14	0.30	0.697
Q × T × L × A	3, 57	7.04	0.003	0.35	0.27	0.608
500-800ms						
T × H × L	1, 19	4.48	0.048	0.49	0.19	n.a.
T × L × A	3, 57	10.54	< 0.001	0.43	0.36	0.647
T × H × L × A	3, 57	7.74	0.001	0.15	0.29	0.705
Q × T × H × A	3, 57	5.05	0.016	0.35	0.21	0.567
800-1300ms						
Q × L × A	3, 57	5.20	0.008	0.53	0.21	0.746
T × L × A	3, 57	5.91	0.005	0.43	0.24	0.772
Q × T × H × A	3, 57	6.31	0.006	0.17	0.25	0.604

Note: Q abbreviates factor of quantifier.

Table 4

Experiment 3 ERP Analysis of Variance

Source	*df*	*F*	*p*	*MSE*	η_p^2	ε_{GG}
Midline electrodes						
300-500ms						
T	1, 23	27.47	< 0.001	11.58	0.54	n.a.
T × A	3, 69	13.14	< 0.001	1.58	0.37	0.602
500-800ms						
T × A	3, 69	7.52	0.003	1.52	0.24	0.582
800-1300ms						
T × A	3, 69	7.23	0.003	1.39	0.24	0.596
Mediolateral electrodes						
300-500ms						
T	1, 23	25.60	< 0.001	35.96	0.53	n.a.
T × L	1, 23	22.99	< 0.001	4.64	0.50	n.a.
T × A	3, 69	7.06	0.010	3.70	0.23	0.403
T × L × A	3, 69	4.28	0.018	0.48	0.16	0.710
500-800ms						
T × A	3, 69	5.14	0.023	2.92	0.18	0.438
T × L × A	3, 69	9.85	< 0.001	0.60	0.30	0.624
800-1300ms						
T × A	3, 69	9.11	0.003	2.77	0.28	0.429