Verb omission errors: Evidence of rational processing of noisy language inputs

Leon Bergen (bergen@mit.edu)[1], **Roger Levy (rlevy@ucsd.edu)**[2], **Edward Gibson (egibson@mit.edu)**[1]
[1]Department of Brain and Cognitive Sciences, MIT, Cambridge MA 02139,
[2]Department of Linguistics, UC San Diego, La Jolla CA 92093

Abstract

We investigate the mechanisms that allow people to successfully understand language given noise in the world and in their own perceptual inputs. We address two parts of this question. First, what knowledge do people use to make sense of language inputs that may have been corrupted? Second, how much of this knowledge is used while people are processing sentences? We conduct a sentence production experiment and an on-line reading experiment in order to answer these questions. Both experiments provide evidence that syntactic knowledge can drive top-down reinterpretations of word identities, as well as syntactic reanalyses that are incompatible with people's language input. In addition, Experiment 2 provides evidence that this knowledge is deployed on-line, as people process sentences.

Keywords: Language Understanding; Sentence Production; Rational Analysis

Introduction

People typically understand language effortlessly and successfully despite the fact that the language input can be corrupted in many ways between the speaker's planning of an utterance and the listener's comprehension of it. This suggests that the language comprehension mechanism has developed methods for correcting noise added to the input, in order to recover the speaker's true intent. We will be viewing this problem from the perspective of rational Bayesian inference. According to this position, people identify the intended language input by rationally integrating their prior linguistic expectations with the possibility of noise. The claim that people are at least somewhat rational when correcting for noise seems obviously true – if someone hears "The teacher write", they are unlikely to perceive this as a totally different phrase, but are instead likely to interpret it as the very similar "The teachers write" or "The teacher writes." A more interesting question therefore is: how thoroughly do rational noisy-channel effects permeate real-time comprehension at different levels and detail of linguistic representation?

This question can be broken into two parts:

1. What kinds of knowledge are deployed during processing in order to determine if an error has occurred in the input?

2. What kinds of reanalyses due to noise does the sentence processing mechanism pursue? Is it restricted to the reanalysis of single words, or will it pursue more radical reanalyses of the syntactic structure?

In order to address these questions, we will first discuss what is entailed by an ideal-observer perspective on sentence processing in the presence of noise. We will then survey some open questions about the extent to which the sentence processing mechanism is adapted to noise. Finally, we will present the results of a sentence production and a sentence comprehension experiment that each bear on these questions.

Noisy channel models

According to a noisy channel model of sentence comprehension, a rational comprehender uses their perceptual input, which may have been corrupted, to identify the language input intended by the speaker. There are two sides to the comprehender's inference problem. First, prior linguistic knowledge should constrain people's inferences about what the speaker actually meant. If a particular phrase is very unlikely *a priori*, then it is even more unlikely that it was intended by the speaker but later corrupted by noise. For example, people's syntactic knowledge will inform how they interpret an ungrammatical sentence like "The voter hope there will be a recount." There is a conflict between the two parts of this sentence: the phrase "The voter hope" is a noun-noun compound, but the phrase "there will be a recount" is the argument of a verb. This conflict can be resolved by a simple syntactic reanalysis of the first phrase. If the first phrase is reanalyzed as a noun-verb construction, then the second phrase will have its required verb, and the sentence will be grammatical. As a result, it seems very compelling to reanalyze the first phrase as "The voter hopes" or "The voters hope."

While people's linguistic knowledge constrains their inferences about the intended input on one side, their knowledge about the process that generated the noise constrains them on the other. People do not usually intend to say one thing and then accidentally say something completely unrelated. Consequently, the noise process makes certain hypotheses about how the input was corrupted very unlikely. Similarly, our perception does not usually introduce massive errors into the input we receive. Hence in the example above, people only consider similar sentences when they are trying to figure out the intended meaning of "The voter hope."

The problem of inferring the true intended input can be posed in terms of optimal Bayesian inference. The comprehender's prior linguistic information can be represented by a probability distribution P_L, which is defined over phrases or sentences. The model of the noise process is given by the distribution P_N. This distribution specifies how likely a particular noise event is to occur, e.g. the deletion of a letter in the perception of the input, or the accidental insertion of a morpheme by the speaker. By Bayes' Theorem, the probability that a sentence or phrase s was intended given the perceived sentence or phrase s' is equal to:

$$P(s|s') = \frac{P_L(s)P_N(s \to s')}{P(s')} \quad (1)$$

where $P_N(s \to s')$ is the probability that s' will be perceived when s is intended by the speaker.

In general, having perceived the input s', we can measure the comprehender's evidence for phrase s_1 relative to s_2 by looking at the ratio

$$\frac{P(s_1|s')}{P(s_2|s')} = \frac{P_L(s_1)P_N(s_1 \to s')}{P_L(s_2)P_N(s_2 \to s')}. \tag{2}$$

The higher this value, the more evidence that s_1 and not s_2 was intended. We can apply this formula to the case that $s_2 = s'$. The resulting ratio

$$\frac{P(s|s')}{P(s'|s')} = \frac{P_L(s)P_N(s \to s')}{P_L(s')P_N(s' \to s')}. \tag{3}$$

can be interpreted as the probability that an error occurred and s was intended relative to the probability that no error occurred and s' was in fact intended.

These formulas capture some important intuitions about how people should infer the intended input, and also have some interesting consequences. Only sentences s that have a relatively high probability of causing the perceptual input s' will be plausible candidates for the intended meaning of the speaker. If this probability $P_N(s \to s')$ is low, then by equation 1, the probability $P(s|s')$ that s was intended must also be low. For example, if sentence s is very different than s', then it would require a large number of specific errors to transform s into s'. As a result, $P_N(s \to s')$ would be low, and there would be low probability that s was actually intended.

The model also captures an interesting tradeoff between the comprehender's linguistic knowledge and the noise model. It will always be the case that the easiest way to generate a perceptual input is for that perceptual input to have been intended: it is not necessary to posit any errors in this case in order to explain the perceptual input. However, the comprehender may infer the presence of noise when there is an alternative sentence s that is sufficiently similar to s' and has higher probability according to the language model P_L. Formula 3 states that as the ratio $\frac{P_L(s)}{P_L(s')}$ of the probability of s relative to s' increases, the comprehender will be more likely to infer that s was intended. One simple consequence of this is that spelling mistakes will be corrected by the comprehender: misspelled words will receive low probability under the model P_L relative to nearby, correctly spelled words.

A more interesting consequence is that under certain conditions, the comprehender should infer that noise was added to the input even when the actual sentence has a legitimate analysis. This can happen when the perceived sentence is well-formed but highly unlikely because of semantic implausibility or because it contains low-frequency syntactic constructions. If there is a sufficiently similar sentence s that has higher probability under the language model, then the ratio in formula 3 may be high enough for the comprehender to infer that s was actually intended.

Finally, the noisy channel model should lead us to predict that comprehenders will treat the presence of a language-unit (e.g. a letter, morpheme, or word) differently from its absence. In particular, comprehenders should be more likely to infer that a perceived language unit was intended by the speaker than that the absence of a language unit was intended.

The Bayesian size principle of (Xu & Tenenbaum, 2007) explains this asymmetry. This principle states that when a hypothesis allows for many possibilities to occur, it must place a small amount of probability on each of these possibilities. More formally, the principle states that

$$P(h|H) = \frac{1}{|H|} \tag{4}$$

where H is a hypothesis containing h, and $|H|$ is the number of possibilities contained in H. In the present setting, this implies that there is a lower chance of any specific language unit being accidentally inserted into a perceived sentence than there is of a specific language unit being accidentally deleted. To see this, imagine that because of perceptual noise, a random letter is either deleted from or inserted into part of the sentence. While there are 26 English letters that could be inserted at a given location, there is only one way to delete a specific letter. It follows that there is a higher probability that a particular sentence s_1 was intended, but a letter was deleted from it, than that another sentence s_2 was intended, but a letter was inserted into it. The same reasoning applies to noise generated from speech errors, or insertions and deletions that occur on larger language units such as morphemes and words.

Previous work

While these types of comprehension behavior are predicted by an ideal-observer model, it is unclear if they are borne out in human comprehension. Previous work has addressed parts of this question. A number of studies have shown the influence of non-syntactic factors on how people determine word identity. In the speech perception literature, researchers have demonstrated context effects in word recognition (Marslen-Wilson, 1975, 1987; Dahan & Tanenhaus, 2004). However, our understanding of two aspects of word recognition remains more limited: (i) the role of syntactic ambiguity as a contextual factor for word recognition; (ii) the consequences of misrecognizing a word, or remaining uncertain as to its identity, for downstream comprehension of the sentence. Event plausibility has also been shown to induce word misrecognition and subsequent processing difficulty, but such processing difficulty has not yet been demonstrated when different syntactic analyses are involved (Slattery, 2009).

Other modeling and experimental studies have provided evidence for the effect of syntax on the correction of noise. Levy (2008) proposed a Bayesian model of noisy-channel sentence comprehension similar to the present one, and suggested that such models might shed light on a number of syntactic-comprehension phenomena difficult to reconcile with traditional sentence-processing models. Gibson and Bergen (2012) provided evidence suggesting that comprehenders can come to global misinterpretations of complete sentences that are incompatible with the true string on the basis of both semantic and syntactic information, but this work does not show how these misinterpretations unfold in real time, and only considers cases in which a word is inserted or swapped, and not those in which a word is substituted. There is also evidence that comprehenders can be induced to disregard orthographic material (commas) in reading on the basis of strongly biased grammatical expectations, and to adopt a garden-path syntactic analysis that should be incompatible with the actual orthographic material (Levy, 2011), but this has not been shown to happen when actual word identities are at stake. Finally, there is evidence that comprehenders can entertain the possibility that a previous word was misrecognized on the basis of alternative syntactic analyses of an input prefix (Levy, Bicknell, Slattery, & Rayner, 2009). However, this does not show that comprehenders will actualy pursue these alternative syntactic analyses further in the sentence.

The present studies aim to fill several gaps in this literature. We will provide evidence that the misrecognition of words can be driven by the comprehender considering grammatical analyses that are incompatible with the true input and taking into account fine-grained (word-word-collocation-dependent) preferences for syntactic analysis. Our evidence will suggest that comprehenders pursue these alternative analyses downstream in the sentence. Finally, we will provide evidence that comprehenders' inferences about noise are rationally sensitive the asymmetry between insertions and deletions.

Experiment 1: Verb omission errors

We discussed two distinctive claims of the noisy channel model in the previous section:

1. Comprehenders should infer that a perceived phrase contains an error when there exists a nearby syntactic construction with a much higher base-rate of occurrence. This is true even when the phrase appears well-formed.

2. Comprehenders should be more likely to infer that part of the intended phrase was deleted than that part was inserted. As a result, they should be more likely to believe that a phrase was intended when the only nearby alternatives could have only produced the phrase via insertion.

We test these claims using a timed sentence completion task, with participants asked to complete short sentence preambles. Participants were briefly shown a sentence preamble. This preamble then disappeared, and they were asked to type a complete sentence beginning with this preamble.

Participants were shown preambles that were unambiguously noun-verb ("NV"), unambiguously noun-noun ("NN"), or NN/NV ambiguous constructions (Frazier & Rayner, 1987; MacDonald, 1993). Crucially, each NV preamble differed by only a single morpheme ("-ed") from an NN preamble, and similarly for all of the NN preambles. For example, "The immigrant feared" was used as an unambiguous NV construction, because "feared" must be interpreted as the main verb of the resulting sentence, while "The immigrant fear" was used as an unambiguous NN construction, because "fear" must be interpreted as a noun in the resulting sentence. In addition, the preambles varied in their degree of bias towards the NV or NN constructions. A summary of these conditions is provided in Table 1. The purpose of the ambiguous-preamble continuations was to validate the categorization of each item as NN-biased or NV-biased (see Results below).[1]

Table 1: Experiment 1 conditions

Condition	Example
NV-biased/NN preamble	The immigrant fear
NV-biased/NV preamble	The immigrant feared
NV-biased/ambiguous preamble	The immigrant fears
NN-biased/NN preamble	The almond roll
NN-biased/NV preamble	The almond rolled
NN-biased/ambiguous preamble	The almond rolls

Our noisy channel model predicts a specific pattern of *rational misidentification* of the unambiguous sentence preambles. Participants should primarily misidentify less likely constructions in favor of more likely constructions, and should primarily infer that a morpheme was dropped from a construction, not that it was added. Under the model, therefore, there are two criteria both of which must hold for misidentifications to be most frequent. The first is that the preamble be NN, since NV can be converted to NN by deletion of a single letter or past-tense morpheme, whereas NN-to-NV conversion requires an insertion. The second is that the preamble be NV-biased, since the NV construction is *a priori* more likely in this case. We thus predict that NV-biased NN preambles should be the most likely to be misidentified; in each of the other conditions at least one of the two above criteria fails to hold, and we therefore expect fewer misidentifications of the preambles in these cases.

There are several ways that such a pattern of rational misidentification could manifest itself in our sentence completion paradigm. First, if participants misidentify the preamble, then we expect *repetition errors* in which they retype the preamble incorrectly. For example, if they misidentify an NN preamble as an NV preamble, then we expect them to retype a preamble containing a verb with a past-tense morpheme.

We may also find a more interesting effect which results from participants' rational *uncertainty* about the identity of the sentence preamble. The evidence that participants receive from an NV-biased NN preamble may lead them to be rationally uncertain about whether an NN or NV preamble was intended, as there is still a tradeoff between positing noise and moving to a more probable construction. An optimal sentence processing mechanism would maintain representations of both of these possibilities after encountering the preamble.

We may find evidence of these multiple representations in the form of *verb omission errors*. In such sentences, participants would correctly reproduce the NN preamble, but complete the sentence as though they had already introduced a

[1] Some of our unambiguous preambles do have alternate syntactically permissable readings, e.g. "The almond rolled ice cream was good," but these are low-frequency constructions, and participants never used them in their completions.

main verb. The following sentence provides an example:

(1) The immigrant fear being deported.

The word "fear" is a noun here, though it would need to be a verb for the sentence to be grammatical. If participants produce such sentences, then this would be evidence that they are maintaining uncertainty about the identity of the preamble, and are repeating the preamble according to one representation, but completing it according to the other.

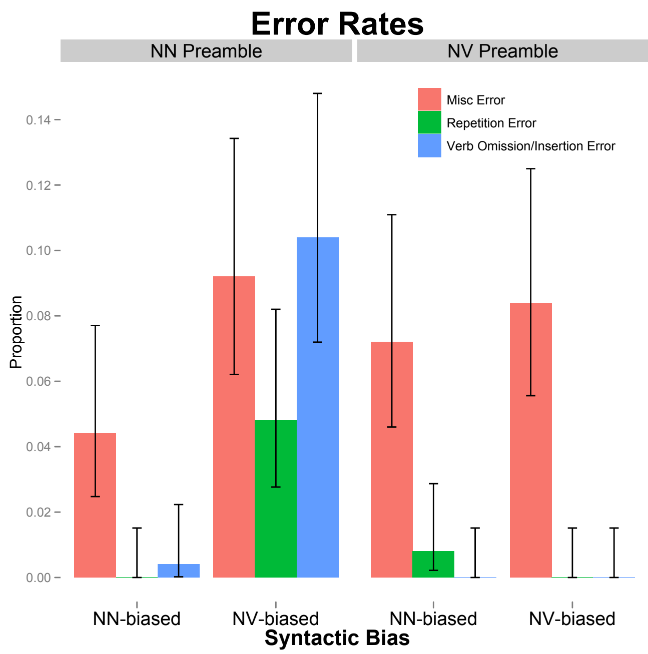

Figure 1: Experiment 1 results. The left panel shows the error rates in the NN condition, depending on the syntactic bias of the construction, while the right panel shows the error rates in the NV condition. Error bars are 95% confidence intervals. Note that, according to the coding criteria, verb omission errors occurred only in the NN-preamble condition, while verb insertion errors occurred only in the NV-preamble condition.

Methods:

Participants: Sixty native English speakers from the United States were recruited from Amazon Mechanical Turk. They were paid a small amount of money for participation.

Materials and Procedure: Participants were shown each sentence preamble for 1.5 seconds. After this, the preamble disappeared from the screen, and participants were given 13 seconds to retype the preamble exactly and complete the sentence. During the instructions, participants were shown several instances of incorrect completions; for example, they were told that if they were asked to complete "The dog in the park", then they could not add anything to the preamble as in "The dog in the parks."

Items consisted of 12 NN-biased and 12 NV-biased preambles that were selected for their bias. These preambles were presented in one of three conditions: NN, NV, or an ambiguous condition. These items were presented in a within-subjects design. The ambiguous condition consisted of plural or present-tense items of the form "The immigrant fears", which is consistent with either an NN or NV reading. This condition was used for norming the items: higher rates of NV completions indicated NV-bias, and similarly for NN completions. Finally, 12 unambiguous NN and 12 unambiguous NV items were used as fillers; these were distinct from the test items in not being in the morphological neighborhood of an alternative construction.

Coding: Completed sentences were coded as correct if the sentence preamble was repeated correctly and the sentence was grammatically well-formed. A sentence was coded as a repetition error if its preamble was repeated with a morphological error that switched its grammatical role (e.g. if an NN preamble was repeated with past-tense morphology) and the rest of the sentence was grammatically consistent with the repeated preamble. In the NN conditions, a sentence was coded as a verb omission error if the preamble was repeated correctly but the rest of the sentence grammatically required a main verb to appear in the preamble. In the NV conditions, a sentence was coded as a verb insertion error if the preamble was repeated correctly but the rest of the sentence grammatically required the preamble to be a noun phrase.

Of the total responses, 7% contained miscellaneous errors, which did not fall under the other criteria. For 74% of these errors, the sentence did not contain a complete independent clause; most of the remainder contained word substitutions in the preamble, or number or tense agreement errors.

Results and discussion

We first analyzed whether the NN and NV-biased items were biased in the correct direction. This was determined using the completions for the ambiguous condition. All 12 of the NN-biased items received NN completions most frequently; 11 of the 12 NV-biased items received NV completions most frequently, and 1 was equi-biased. Of the 229 NN-biased items in the ambiguous condition completed correctly, 197 (86%) were given NN completions. Of the 228 NV-biased items completed correctly, 184 (81%) were given NV completions.

We next looked at whether repetition errors and verb omission errors occurred more frequently for the NN preambles. We tested this using repeated measures ANOVAs; Figure 1 shows the frequencies of each type of error for each experimental condition.[2] There were significantly more repetition errors in the NN preambles than in the NV preambles (12 vs. 2; $p < 0.05$). In addition, there were significantly more verb omission errors in the NN preambles than noun omission errors in the NV preambles (27 vs. 0; $p<0.001$). We note that for the NN preambles, the repetition errors were approx-

[2] We did not use mixed logit models because of convergence issues with random slopes.

imately evenly distributed among errors to the plural marking of the first noun, past-tense errors on the second noun, and present-tense errors on the second noun.

Our final question was whether the error rate for the NN items was higher for the NV-biased items (we did not analyze this for the NV items, because these contained so few errors). Because the bias was manipulated between-items, we tested this using ANOVAs with subjects as random factors. All of the repetition errors for the NN items occurred on NV-biased items (12 vs. 0; $p < 0.01$). All but one of the verb omission errors occurred on NV-biased items (26 vs. 1; $p<0.001$).

These results provide evidence for both predictions of the noisy channel model. First, we observed the predicted asymmetry between insertions and deletions: nearly all of the observed errors were consistent with the participants inferring that a morpheme had been dropped from the preamble they observed. Second, we observed that nearly all errors were made in the direction of the more probable construction.

We also found evidence for a more tentative prediction of the model, which is that participants would maintain uncertainty about the identity of the preamble. Verb omission errors patterned in the same manner as the repetition errors; moreover, they occurred at twice the rate of the repetition errors. This suggests that participants were maintaining multiple representations of the sentence preamble; verb omission errors occurred when more than one representation was deployed in the completion of the sentence.

Experiment 2

In Experiment 2, we used a different method to evaluate the predictions of the noisy channel model. The model predicts that people will adopt incorrect syntactic analyses if there exist similar phrases that could have easily generated them. In such cases, we should be able to observe the effects of misidentification downstream in sentence comprehension: specifically, comprehenders should have difficulty if the later parts of a sentence are inconsistent with their interpretation.

We used NN and NV preambles like those in Experiment 1 to probe such garden-path effects. If people misidentify an NN preamble as an NV preamble, then they should be surprised when a main verb is used later in the sentence; conversely, if they misidentify an NV preamble as an NN preamble, then they should be surprised when the clause ends without a main verb. For example, in the Dense-neighborhood/NN condition in Table 2 – which is so named because there are other grammatical phrases in the morphological neighborhood of its preamble – people should sometimes infer that the past tense "chauffeured" was intended instead of "chauffeur", and therefore they should be surprised when they arrive at the main verb "hoped." On the other hand, if they infer that "chauffeur" was intended instead of "chauffeured" in the Dense-neighborhood/NV condition, then we would expect difficulty at "but", which indicates the end of the first clause. Such effects have been shown for truly ambiguous NN/NV preambles such as "The voter hopes" (Frazier & Rayner, 1987), and have also been shown to be affected by the collocation's lexical bias (MacDonald, 1993), but have never been demonstrated when the preamble is unambiguous.

We used self-paced reading to evaluate whether people had difficulty at these "disambiguating" regions (note, however, that these items are only ambiguous given the possibility of noise). People's performance at these regions was evaluated against two control conditions, Sparse-neighborhood/NN and Sparse-neighborhood/NV, which were given this name because their preambles were not in the morphological neighborhood of alternative syntactic constructions. These conditions were identical to the Dense-neighborhood conditions, except for this difference in the density of the morphological neighborhood. For the Sparse-neighborhood/NN condition, this was done by using an adjective before the head noun, and for the NV condition, this was done by using the quantifier "some" and a plural morpheme before the main verb.

The noisy channel model predicts an interaction between the density of the morphological neighborhood and grammatical structure at the disambiguating region. In particular, because of the asymmetry between morpheme insertion and deletion, people misidentify the preamble most frequently in the Dense-neighborhood/NN condition. We should expect less difficulty at the disambiguating region of the Dense-neighborhood/NV condition because the syntactic alternative of the preamble would require an insertion to produce the perceived wordform. We should similarly expect less difficulty in both Sparse-neighborhood conditions, because these preambles are far from alternative syntactic constructions.

Table 2: Experiment 2 conditions

Condition	Example
Dense-neighborhood/NN	The intern chauffeur for the governor hoped for more interesting work.
Dense-neighborhood/NV	The intern chauffeured for the governor but hoped for more interesting work.
Sparse-neighborhood/NN	The inexperienced chauffeur for the governor hoped for more interesting work.
Sparse-neighborhood/NV	Some interns chauffeured for the governor but hoped for more interesting work.

Methods:

Participants: We recruited 120 native English speakers from the United States from Amazon Mechanical Turk. They were paid a small amount of money for participation.

Materials and Procedure: We tested participants using a self-paced reading program that ran in participants' web browsers. Words were presented one at a time. After each sentence, participants were asked a comprehension question.

We constructed 16 items in the conditions shown in Table 2. These conditions varied in a within-subjects design. We also included 32 filler sentences during testing.

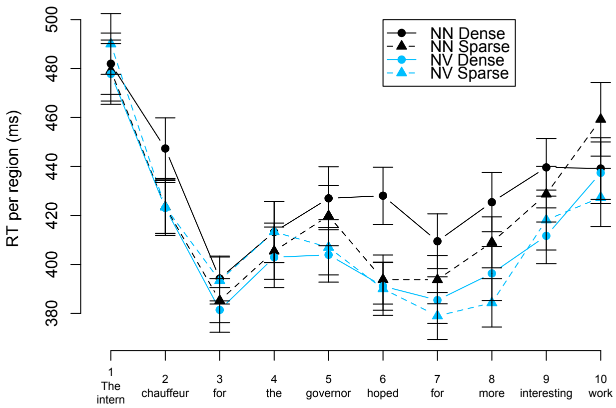

Figure 2: Reading-times from Experiment 2. Region 6 is the critical disambiguating region.

Results and discussion

Before analyzing the data, we excluded participants that answered fewer than 80% of the comprehension questions correctly, and excluded trials on which reading times were farther than 2 standard deviations from the mean. Accuracies on comprehension questions were above 90% on all conditions. The results are shown in Figure 2. For the analysis, words were aligned relative to the disambiguating word, which is labelled as region 6 and is the first place we can expect to find any critical effects of disambiguation.

Our first question was whether there was an effect of neighborhood density for the NN items, for which we expected difficulty at the disambiguating region. We performed out analysis with a linear mixed-effects model with random slopes and interactions for participants and items. We found a significant increase in RTs at the disambiguating region for the Dense-neighborhood items (β=33.08 ms, t=2.49, $p < 0.05$). We next looked at whether this effect was strongest for the NN items. There was a significant interaction between neighborhood density and syntactic structure (β=35.06 ms, t=2.21, $p < 0.05$), indicating a superadditive effect of density and structure on RTs. No interactions were significant ($p>0.05$) at any region prior to the critical region.

These results provide evidence that participants misidentified the sentence preambles, and that this misidentification was rational. Specifically, we found evidence that participants were most likely to misidentify preambles that could have been produced by an alternative phrase via a deletion. This is consistent with the asymmetry between insertions and deletions implied by the Bayesian size principle.

Discussion

We have investigated the knowledge that people use to correct noise in their language input, as well as the on-line processing mechanisms that support this error correction. Regarding the first topic, our experiments provide evidence that comprehenders entertain and even adopt syntactic reanalyses of their language input to account for the possibility of noise, even when these reanalyses are inconsistent with the true input. Moreover, we found evidence that these reanalyses are driven by the rational integration of grammatical expectations and expectations about the noise process. The results of Experiment 1 suggest that people prefer alternative explanations for their input that involve higher probability syntactic constructions, and noise consisting deletions rather than insertions.

We have also found evidence that people employ these alternative syntactic analyses during on-line sentence processing. In Experiment 2, we found downstream effects of noise-driven reinterpretations at the point when they contradicted the input sentence. This suggests that the sentence processing mechanism is not delayed in positing or making use of syntactic reanalyses during the course of comprehension. This is the first result demonstrating that comprehenders will pursue garden-path syntactic analyses differing from the true sentence preamble by a word substitution, and that garden-path disambiguation in these cases incurs measurable costs; this result is directly predicted by our noisy-channel model. Together with other recent work, these results raises new questions regarding the full breadth of sentence-processing phenomena that may be best understood as the consequence of rational, noisy-channel probabilistic inference.

References

Dahan, D., & Tanenhaus, M. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 498.

Frazier, L., & Rayner, K. (1987). Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. *Journal of memory and language*, *26*(5), 505–526.

Gibson, E., & Bergen, L. (2012). The rational integration of noise and prior semantic expectation: Evidence for a noisy-channel model of sentence interpretation. *Submitted*.

Levy, R. (2008). A noisy-channel model of rational human sentence comprehension under uncertain input. In *Proceedings of the 13th conference on empirical methods in natural language processing* (pp. 234–243).

Levy, R. (2011). Integrating surprisal and uncertain-input models in online sentence comprehension: formal techniques and empirical results. In *Proceedings of the 49th annual meeting of the association for computational linguistics.*

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences*, *106*(50), 21086.

MacDonald, M. (1993). The interaction of lexical and syntactic ambiguity. *Journal of Memory and Language*, *32*, 692–692.

Marslen-Wilson, W. (1975). Sentence perception as an interactive parallel process. *Science*, *189*(4198), 226.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1), 71–102.

Slattery, T. (2009). Word misperception, the neighbor frequency effect, and the role of sentence context: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1969.