Ungrammatical Influences: Evidence for Dynamical Language Processing

Whitney Tabor
tabor@uconnvm.uconn.edu
Bruno Galantucci
Bruno.Galantucci@uconn.edu

Department of Psychology
University of Connecticut
Storrs, CT 06269 USA

Abstract

A distinguishing feature of self-organizing models of cognitive structure is that they permit incompatible structures to coexist at least temporarily. Here we report on a connectionist model of natural language processing which appears to temporarily construct incoherent structures. We then describe two reading-time studies which reveal people exhibiting the same tendency. In particular, both networks and people show sensitivity to the irrelevant structural interpretations of the underlined phrases in (1) and (2).

(1) We did not think the company would fire truck drivers without consulting the union first.
(2) The manager watched the waiter served pea soup by the trainee.

This kind of sensitivity is absent in parsing models which treat grammatical constraints as absolute because such models lack a principled method of generating incoherent parses. Connectionist networks make the right predictions by using feedback and self-organization. Our results push in the direction of seeking a solution to the tractability problems of parsing by using dynamical mechanisms in a parallel architecture.

Introduction

Current sentence-processing research tends to focus on ambiguity-related processing in sentences like (1) – (3):

(1) The mechanic maintained the truck was working beautifully.

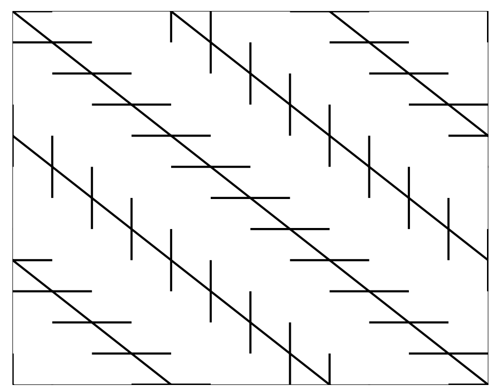
(2) The cop arrested by the detective was chagrined.

(3) The cook stirred the soup with the tomatoes.

Each of these sentences has a structural ambiguity which is resolved on the basis of structural or pragmatic information when the underlined words arrive. Reading time and eye tracking studies show that when biases favor the wrong interpretation initially, readers tend to slow down and/or make regressions in the disambiguating region, which suggests that they either choose the wrong parse initially or are biased toward it (see Frazier, 1988; Tanenhaus & Trueswell, 1995).

Such phenomena accord well with a model of sentence processing which assumes people construct phrase-structure parses incrementally based on the input up to the current point in time. On this view, the slow-down in the disambiguating region is due either to extra time spent on revising an incorrect parse, or to extra time spent on revising the weighting assigned to different possible parses maintained in parallel.

Figure 1: The Zöllner illusion.



Focusing for a moment on cognitive processes outside of sentence processing, there is a good deal of evidence that people are reliably vulnerable to certain adverse influences when interpreting complex stimuli. In the Zöllner illusion (Held, 1971—Figure 1), lines on a page appear to be nonparallel even though retinal and depth of field information indicate parallelism. Similarly, in the Stroop effect (Stroop, 1966), a decision is supposed to have been made ahead of time to interpret the stimulus along one particular dimension of contrast (e.g. color), and yet when the stimulus is presented, people are often led astray by irrelevant verbal information.

These cases are different from the classic sentence processing examples listed in (1) through (3) in that they show people temporarily failing to rule out an interpretation that could be ruled out absolutely, given the information at hand. What would be the analogous cases in sentence processing?

Definition of Ungrammatical Influences

There is a class of sentences in which one parse of a word sequence can be completely ruled out on grammatical grounds and yet (we hypothesize) people are influenced by it anyway. The following are examples of such hypothesized "Ungrammatical Influences":

(4) a. They won't fire truck drivers on Sunday.
b. They won't hire truck drivers on Sunday.

(5) a. The manager watched the waiter served pea soup by the trainee.
b. The manager watched the waiter given pea soup by the trainee.

Each of the (a) examples has a familiar construction within it that is irrelevant to the only grammatical parse of the sentence. But by the time this distractor construction is encountered, it can be ruled out on grammatical grounds. Our hypothesis is that people are influenced by this "ruled out" parse nevertheless. Thus the (a) examples should be processed differently from the (b) examples which lack the distractors. In (4), the sequence of words "fire truck" forms a familiar compound in English, but coming on the heels of a modal verb, "would", the word "fire" can only reasonably be interpreted as a verb, not a noun. Similarly, in (5a), the second verb "served" must be interpreted as a passive verb introducing a reduced relative clause which modifies the noun phrase, "the waiter". But, taken in isolation, "the waiter served pea soup" makes a sensible transitive construction with an active verb.

Our hypothesis is that readers will be distracted by these pockets of coherent structure, even though the structures are incompatible with prior information.

Models

We find that an often-studied connectionist network, the Simple Recurrent Network (or SRN), behaves in accordance with the hypothesis that Ungrammatical Influences exist. This prediction distinguishes it from most current models of sentence processing.

Elman (1991) showed that a recurrent connectionist network trained by and approximation of backpropagation through time (Rumelhart, Hinton, and Williams, 1986) on word prediction could extract much of the structure of a natural-language-like generating process from a corpus generated by the process.

We trained such a network on the output of Grammar 1 (see Table 1). The network was trained on the task of predicting next words in a constantly growing corpus of strings generated by Grammar 1. The sentences were presented to the network one word at a time. Each input unit corresponded to a possible current word and each output unit corresponded to a possible next word (Elman, 1990, 1991). The learning rate was set to 0.01 throughout and no momentum was used.

The network's output layer had normalized exponential units. During training, error on a given word was thus defined as the Kullback-Leibler Divergence between the vector of network output activations and the output encoding of the next word that occurred in the corpus (Rumelhart, Durbin, & Chauvin, 1995). We stopped training when the network had successfully distinguished the underlying states of the grammar. At this point, it had seen on average about 500,000 words in sequence.

Since optimal training of such networks causes the output activations to converge on the expected value of the outputs given the inputs, we computed the Kullback-Leibler Divergence between the output activation pattern and grammar-derived probability distributions for each string of interest. The average Divergences over the six test and control sentences of the form (4) from the grammar are shown in Figure 2.

Table 1: A simple phrase structure grammar for generating Noun Noun compounds and Noun/Verb ambiguities.

0.50 S → SVP
0.50 S → SNP

0.17 SVP → to waste N[Obj] is unforgivable
0.17 SVP → to bear N[Obj] is necessary
0.17 SVP → to mail N[Obj] is costly
0.17 SVP → to place N[Obj] is challenging
0.16 SVP → to cart N[Obj] is toilsome
0.16 SVP → to fuel N[Obj] is ignoble

0.17 SNP → the waste baskets are large
0.17 SNP → the bear cubs are round
0.17 SNP → the mail men are persistent
0.17 SNP → the place mats are flat
0.17 SNP → the cart wheels are shaky
0.16 SNP → the fuel tanks are full

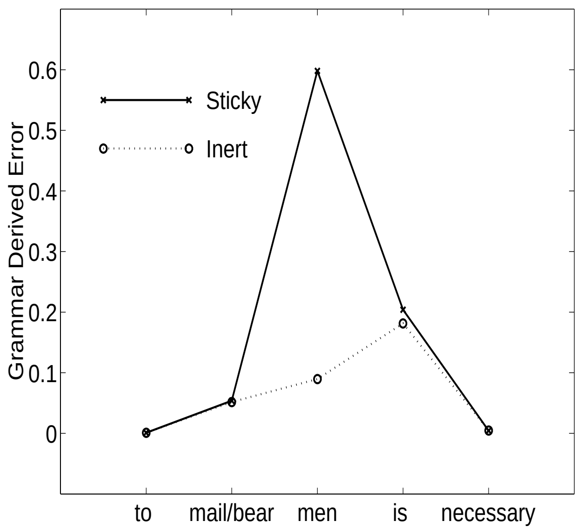
NObj → baskets, mats, cubs, wheels, tanks, men

We repeated the simulation on 10 networks that started learning with different random initial weights. The contrast between the Sticky and Inert conditions occurred in every case. In every case, if we had stopped training earlier (before the network sorted out the differences between states of the underlying grammar), the effect would have been even more pronounced: that is, the network was overwhelmed by the local coherence of the Sticky cases, initially failing to recognize when they occurred in the infinitive context. The effect was somewhat unstable if we trained the network longer on the same materials, and sometimes reversed itself. We suspect that this instability might be reduced if the distractor compounds were not such a prominent feature of the grammar. In real language corpora, coincidences of the Sticky type appear to be quite rare. There are no instances, in the million word Brown Corpus, of coincidental juxtaposition of the 20 sticky pairs used in Experiment 1.

Following Juliano & Tanenhaus (1994), we make an analogy between the network's error scores and reading times in the self paced reading task (Just, Wooley, & Carpenter, 1982) that is often used to study human sentence processing. The network model thus predicts that readers can be distracted by irrelevant interpretations of pairs of words, and that this distraction will lead to higher reading times on the distracting items.

It appears that the Simple Recurrent Network is prone to be distracted by Ungrammatical Influences. By contrast, standard models of syntactic processing, which assume incremental construction of phrase-structure parses, do not predict such effects, for such models insist on global coherence of each parse they construct. There

Figure 2: Simulation 1: Divergence from grammar-derived expected values. Sticky sentences contain irrelevant Noun-Noun compounds immediately after the main verb. Inert sentences do not.



is one class of hybrid Connectionist-Symbolic models which may, with some modifications, predict Ungrammatical Influence effects: it is the class consisting of the the Competitive Attachment Processor ("CAPERS") of Stevenson (1994) and the Dynamical Unification-Space parser of Vosse and Kempen (1999). These parsers build phrase structure trees by positing variable-strength bonds between nodes in a phrase-marker, and allowing incompatible attachment possibilities to compete with each other under a set of constraints which favor globally coherent structures. Both of these frameworks currently assume that words are brought into the "Unification Space" one at a time, and that some resolution is reached before additional words are incorporated. Thus they do not permit local coherences between successive words to give rise to detached substructures. Nevertheless, it is natural to consider the possibility of allowing them to do so. If one were to permit arbitrary local bonding, then these dynamical structure-building models would probably (modulo the setting of some noise and decay parameters) predict Ungrammatical Influence effects.

What, then, is at stake when we ask the question if Ungrammatical Influences exist? Distinguishing properties of the SRN and the hybrid connectionist models are the use of dynamical (continuously adjusting) feedback and self-organization. These models contrast with chart parsers, pushdown automata and other incremental symbolic parsing systems which maximize the use of constraining information at each point in time. Research on incremental symbolic parsing has strained to grapple with tractability problems associated with the combinatorial growth of parse structures. It seems, at first blush, that opening the door to the inclusion of local coherences, as the Ungrammatical Influences hypothesis suggests, will only make matters worse. But this impression may be misleading. The coincident emphasis on feedback mechanisms, which allow efficient elimination of incoherent parses through competition, may be just what is needed to permit a parallel processing solution to the tractability problem. Thus, the significance of finding empirical evidence for Ungrammatical Influences is that it would push us in the direction of seeking such a solution.

We turn, now, to empirical investigation of the hypothesis.

Experiment 1

Tabor and Richardson (1999) compared examples like those in (4a-b) above.

Method

Subjects

Thirty-two undergraduates from Cornell University participated in the experiment. All were native speakers of English. The subjects received course credit for their participation. The experiment lasted for about 30 minutes. The data from one subject was removed from the analysis because of a corrupted file problem.

Materials.

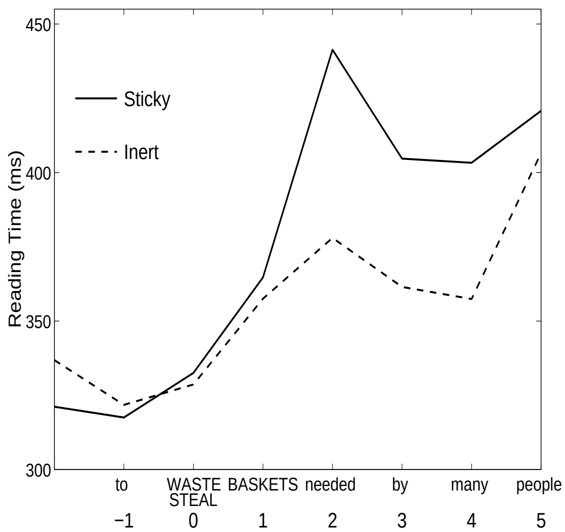
Sixteen target sentences and 16 controls were created. Each target sentence included a clause beginning with a syntactic pattern that strongly constrained the next word to be a verb (e.g., *Some people cannot...* (NP Aux: 7 stimuli), *We decided to...* (NP V[inf] to: 7 stimuli), *on a proposal to...* (P NP to: 1 stimulus), *need a truck to...* (V NP to: 1 stimulus)). This next word (labeled "Word 0" in Figure 3) was lexically ambiguous between a verb sense and a noun sense. In its verb sense, it fit naturally with the preceding and following sentential context, both syntactically and semantically. In its noun sense, this word formed a compound with the word after it ("Word 1"), but this compound did not fit the surrounding context either syntactically or semantically. In 15 of the 16 cases, the compound was a Noun-Noun compound. In one case ("fail safe") the compound was an Noun-Adjective compound. The control sentences were exactly the same as the target sentences except that Word 0 did not form a familiar compound with Word 1. In 14 out of the 16 controls, Word 0 was ambiguous between a verb sense and a noun sense (the two exceptions were "attend" and "flunk"). This control ambiguity was important for ruling out the possibility that any contrast we might observe between target and control sentences might be due to contrasting ambiguity in Word 0.

Procedure.

The sentences were presented using the moving-window self-paced-reading method of Just, Carpenter, and Wooley (1982). Readers read sentences one word at a time, pushing a spacebar to see each successive word. Reading times are measured as intervals between spacebar presses.

The 20 targets and controls were sampled randomly and distributed among 80 filler items. The experiment was preceded by a sequence of six practice trials.

Figure 3: Graph of mean reading time versus position for Experiment 1.



Results. All subjects scored better than 80% on the comprehension questions.

We computed the base 10 logarithms of the raw reading times to normalize their distribution. We performed a linear regression with characters-per-word as independent variable and subjects as random factor. The analyses we report below were performed on the standardized residuals from this regression analysis (Trueswell, Tanenhaus, & Garnsey, 1994).

Figure 3 shows average self-paced reading times at word positions -2 through 5. For each region of interest, subject and item means were subjected to separate analyses of variance (ANOVAs), each with a single factor: Stickiness. The means were not significantly different across the two conditions at any word prior to Word 2 or beyond Word 4. The effect of stickiness was significant in both subject and item analyses in the region defined by Words 2, 3, and 4 together ($F1(1, 30) = 10.77$, $p < .005$; $F2(1, 15) = 4.79$, $p < .05$). The stickiness effect was also significant in the subject analysis at Word 2 alone ($F1(1, 30) = 5.82$, $p < .05$), at Word 3 alone ($F1(1, 30) = 8.78$) $p < 0.01$), and at Word 4 alone ($F1(1, 30) = 6.38$, $p < .05$). Stickiness was marginally significant in the item analysis at Word 3 alone ($F2(1, 15) = 4.35$, $p = .054$) and at Word 4 alone ($F2(1, 15) = 3.51$, $p = .08$).

Discussion.

These results support the claim that Ungrammatical Influences involving two word sequences exist.

But there is an alternative explanation of the outcome should be considered. An early indication of the existence of Ungrammatical Influences came from a priming experiment on the modularity of the lexicon. Tanenhaus, Leiman, & Seidenberg (1979) found that even the irrelevant meaning of a syntactically ambiguous word (e.g. "rose") would cause priming for a short interval (< 200ms) after the word was read in a syntactically constraining context (e.g., "They all rose."). These results are naturally accounted for in a model that assumes that an activation based lexicon operates partially independently of a phrase-building parser. An ambiguous word activates nodes corresponding to all its senses in the lexicon, and irrelevant nodes are only clamped down when syntactic information is later brought to bear. The results of Experiment 1 may reflect such lexical "automaticity", since the two-word locally coherent structures are Noun-Noun compounds, which are arguably lexical items (e.g., Mohanan, 1986). Perhaps the parser correctly chooses to treat these sequences as Noun-Verb collocations, but activation of the compound sense in the lexicon creates interference which slows reading down.

Thus Experiment 1 does not decisively demonstrate the existence of Ungrammatical Influences. The next experiment is designed to probe for the existence of Ungrammatical Influences in a case that does not conform to the lexical activation model's predictions.

Experiment 2: English clauses

Experiment 2

The examples in (5a) contain a potentially distracting local coherence in the form of a clause. It is less convincing that clauses are stored as lexical units since they occur in so many combinations and their meanings can generally be computed compositionally.

Method.

Subjects.

47 subjects were recruited from classes and through advertisement on the campus of the University of Connecticut. All were native speakers of English. They received either money or course credit for their participation. The experiment lasted for about 30 minutes.

Materials.

Eighteen experimental items were created. Each item involved four conditions as in (6):

(6) The manager watched the waiter...

			0	1	2	3	
a.			served	pea	soup	by...	(R / H)
b.	who	was	served	pea	soup	by...	(UR / H)
c.			given	pea	soup	by ...	(R / NH)
b.	who	was	given	pea	soup	by ...	(UR / NH)

Each item included a noun phrase in a non-subject position which was modified by a relative clause in passive voice. Two dimensions of contrast in the relative clause gave rise to four conditions for each item. The relative clause was either reduced (R) or unreduced (UR); its past participle verb was either homophonous and homographic (H) with the corresponding past tense form or distinct from it (NH). Relative clauses like these have been extensively studied in the case where they occur as modifiers of nouns in subject position in a finite clause as in (7) (e.g., Ferreira and Clifton, 1986; Trueswell, Tanenhaus, and Garnsey, 1994).

(7) The waiter served pea soup by the trainee ate ravenously.

Figure 4: Reading times in the four conditions of Experiment 2.

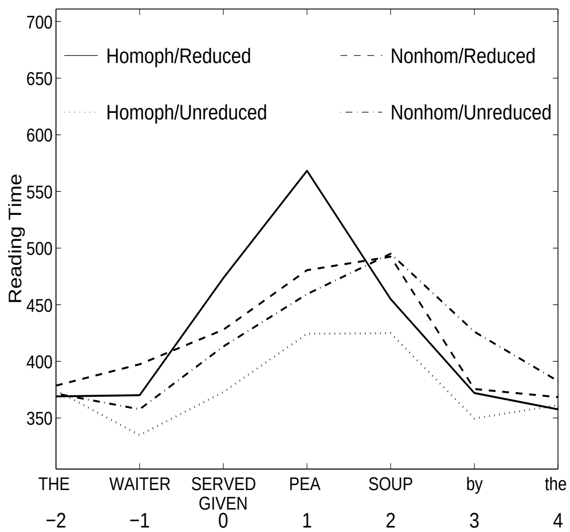
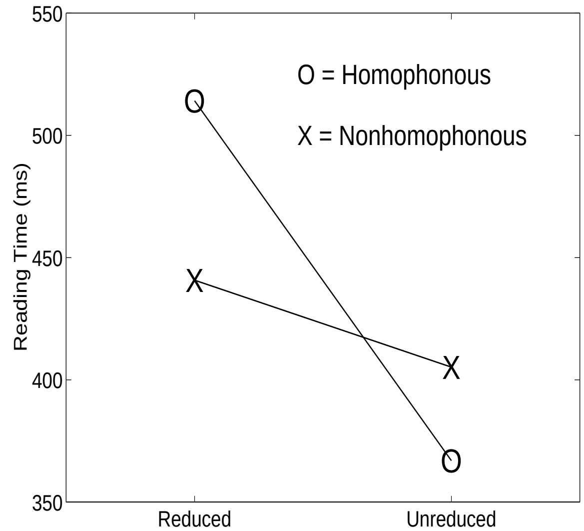


Figure 5: Interaction between Homophony and Reduction in Experiment 2 (Words 0 to 2).



The evidence indicates that when it is semantically sensible to interpret the verb following the subject noun as the main verb of the clause, readers have a strong tendency to do so. Consequently, they become confused starting around the words "by the trainee ate" in a case like (7) because these words disambiguate in favor of the relative clause reading. In a case like (6a), however, the syntax of the words prior to the reduced relative clause precludes the possibility of a main verb reading of the relative clause verb ("served"). If readers were to compute such a reading, then, this would be a case of an Ungrammatical Influence.

We are looking for an effect of Reduction in the Homophonous case. If this effect obtains and the Unreduced cases are read faster than the Reduced cases, the Ungrammatical Influences hypothesis will not be contradicted. However, it would be premature to take such a result on its own as evidence for the existence of Ungrammatical Influences. Greater speed of processing is expected at the relative clause verb in (b) simply because the syntax is more constraining at this point in case (b) than case (a). That is, it is generally the case that processing speed is faster at grammatical events that are more expected (Jurafsky, 1996; Tabor, Juliano, and Tanenhaus, 1997). Thus, we expect a slowing effect of Reduction in the Nonhomophonous case as well ((d) vs. (c)). For this reason, we have employed the more complex 2 x 2 design. We expect that reduction will slow processing in both cases (a) and (c), but it will slow it more in (a) than in (c). If this interaction occurs, then we will have convincing evidence of the existence of Ungrammatical Influences.

Procedure

The procedure was the same as for Experiment 1.

Results. All subjects scored better than 80% correct on the comprehension questions and all the data were used in the analysis.

For each region of interest, subject and item means were subjected to separate analyses of variance (ANOVAs), each with two factors: Homophony and Reduction. There was a main effect of Reduction in the region defined by Words 0 to 2 ($F1(1, 46) = 16.83$, $p < .001$; $F2(1, 17) = 7.71$, $p = .013$). There was a main effect of Homophony in the region defined by Words 2 to 3, that was significant in the subject analysis only ($F1(1, 46) = 21.10$, $p < .001$). In both subject and item analyses, there was a significant interaction between Homophony and Reduction over the region defined by words 0 to 2 ($F1(1, 1) = 26.83$, $p < .001$; $F2(1, 1) = 6.99$, $p = .018$). The interaction was also significant at Word 0 ($F1(1, 46) = 12.31$, $p = .001$; $F2(1, 17) = 11.66$, $p = .004$), and significant by subject at Word 1 ($F1(1, 46) = 6.03$, $p = .018$; $F2(1, 17) = 3.85$, $p = .069$) and Word 2 ($F1(1, 46) = 4.25$, $p = .045$). Figure 4 is a graph of reading times for Experiment 2. Figure 5 is graph of the interaction. As Figure 5 indicates, Reduction slowed reading times in both the Homophonous and the Nonhomophonous conditions, but the slowing was significantly greater in the homophonous case.

Discussion

The existence of the interaction, with Reduction slowing the Homophonous case more than the Nonhomophonous case, supports the Ungrammatical Influences hypothesis.

There is one aspect of the outcome for which we do not have a clear explanation. The distracting effect of the local structural ambiguity affects reading times earlier in Experiment 2 than in Experiment 1, relative to the locally ambiguous region. We speculate that this difference in timing stems from the fairly unusual syntax of the grammaticality correct interpretation of the Experiment 2 sentences. Reduced relative structures with ditransitive verbs are especially unusual, so readers may be working hard to interpret the sentences in the first place, and an additional distraction from an Ungrammatical Influence may easily disrupt processing. By contrast, the syntactic structures of Experiment 1 are very common modal+Infinitive or "to"+Infinitive collocations, so readers may not detect the distracting influence until it has had more time to "sink in". This interpretation again supports a dynamical treatment of information in parsing: some information takes longer to emerge than other information.

Conclusion

We have focused on the hypothesized phenomenon of Ungrammatical Influences: the syntactic parser is expected to be influenced by local, phrasal coherences that are incompatible with the structure of preceding syntactic material. Two experiments supported the existence of Ungrammatical Influences in parsing. Such effects push the theory of parsing strongly in the direction of dynamical, self-organizing models: Ungrammatical Influences occur because the parser is letting all local coherences among words compete to combine into a maximally coherent structure, rather than deductively eliminating parses based on top-down well-formedness constraints.

Although the present experiments suggest treating Ungrammatical Influences as a kind of interference effect (consistent with the class of Limited Resource models of parsing). Ungrammatical Influences may not always get in the way of parsing. Galantucci, Flores D'Arcais, and Tabor (1999) found that when sentences required people to establish reference for a pronoun, and there was a natural candidate embedded in the internal structure of a compound word (e.g., *The killjoy*$_i$ *did not manage to kill it*$_i$ *after all.*), processing was facilitated, even though grammatically, the binding is disallowed. These results, combined with the results discussed in this paper suggest that the theory of grammar needs to take up in earnest the problem of incoherent structure representation.

Acknowledgments

Daniel Richardson was a major contributor to the early stages of this work. Many thanks also to David Perkowski and Kate Finerty who helped with the design and running of the experiments. Helpful comments were provided by Ted Gibson, Neal Pearlmutter and the members of Pearlmutter's lab group at Northeastern University as well as by the members of the University of Connecticut Linguistics/Psychology lunch talk group.

Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning 7*: 295–225.

Frazier, L. (1988) Sentence Processing: A Tutorial Review. In M. Coltheart (Ed.), *Attention and Performance* (pp. 559-586). Hillsdale, NJ: Lawrence Erlbaum Associates.

Galantucci, B., Flores d'Arcais, G.B., Tabor, W. (1999). Italian V+N compounds: evidence for syntactic processing. Presentation at the XIX Conference on Neuropsychology, Brixen, Italy.

Held, R. (1974). *Image, object, and illusion, readings from Scientific American.* San Francisco: W. H. Freeman.

Juliano, C. & Tanenhaus, M.K. (1994). A constraint-based lexicalist account of the subject-object attachment preference. *Journal of Psycholinguistic Research*, 23(6): 459–471.

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science 20*: 137-194.

Just, M.A., Carpenter, P.A., & Wooley, J.D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111, 228-238.

Mohanan, K.P. (1986). *The theory of lexical phonology.* Boston: D. Reidel Pub. Co.

Rumelhart, D. E., Durbin, R., Golden, R., & Chauvin, Y. (1995). Backpropagation: The basic theory. In D.E. Rumelhart & Y. Chauvin, eds. *Backpropagation: Theory, Architectures, and Applications.* Lawrence Erlbaum Associates, Hillsdale, NJ.

Rumelhart, D.E., Hinton, G.E., & Williams, R.J. (1986). Learning Internal Representations by Error Propagation. *In Parallel Distributed Processing, Volume I* (pp. 318-362). MIT Press, Cambridge, Massachusetts.

Stevenson, S. (1994). Competition and recency in a hybrid network model of syntactic disambiguation. *Journal of Psycholinguistic Research*, 23(4):295–322.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18: 643-662.

Tabor, W. & Richardson, D. (1999). Ungrammatical influences in sentence processing. Poster session presented at the 12th Annual CUNY Sentence Processing Conference, New York, NY.

Tanenhaus, M.K., Leiman, J.M., & Seidenberg, M.S. (1979). Evidence for multiple stages in the processing of ambiguous words in syntactic contexts. *Journal of Verbal Learning and Verbal Behavior*, 18:427–440.

Tanenhaus, M. K. & Trueswell, J. C. (1995). Sentence comprehension. In Miller, J. & Eimas, P., (Eds.) *Handbook of Perception and Cognition: Volume 11* (pp. 217–262). Academic Press, San Diego.

Trueswell, J.C., Tanenhaus, M.K. & Garnsey, S.M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language, 33,* 285-318.

Vosse, T. & Kempen, G. (1999). Syntactic structure assembly in human parsing. Poster session presented at the 12th Annual CUNY Sentence Processing Conference, New York, NY.