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

Tabor, W. & Galantucci, B. (2000). Ungrammatical influences: evidence for dynamical language processing. In L. Gleitman & A. Joshi (Eds.), Proceedings of the 22nd Annual Meeting of the Cognitive Science Society (pp. 505–510), Mahhwah, New Jersey: Lawrence Erlbaum Associates.

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 they lack a principled method of generating incoherent parses. Connectionist networks make the right predictions by using feedback and self-organization. We argue that 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 chagrinned.

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

Each of these sentences has a structural ambiguity which is resolved in favor of one structure when the underlined words arrive. Reading time and eye tracking studies show that when pragmatic and/or structural 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.

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 Zolner illusion (Held, 1971), 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 and the task is to respond verbally, people are often lead astray by the 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 analagous 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 distracted by it anyway. The following are examples of such hypothesized "Ungrammatical Influences":

(4) Surely they wouldn't...

a.	...fire	truck	drivers	before	Saturday.
	0	1	2	3	4
b.	...hire	truck	drivers	before	Saturday.

(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 pay attention to this "ruled out" parse nevertheless. Thus the (a) examples should be processed differently from the (b) examples which lack the distractors.

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 NObj is unforgivable
0.17 SVP → to bear NObj is necessary
0.17 SVP → to mail NObj 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

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 in the sentences they read, even though the locally coherent structures are incompatible with prior information.

Models

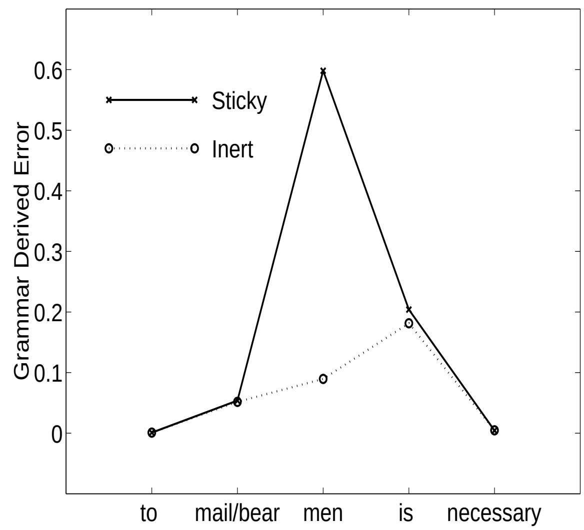
We find that an often-studied connectionist network, the Simple Recurrent Network (or SRN), behaves in accordance with the Don Quixote Hypothesis, and argue that this prediction distinguishes it from most current models of sentence processing.

Elman (1990) and (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 a simple corpus generated by 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 strung end to end and 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.

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



Because the network was predicting probabilities of next words, its output units 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; Rohde & Plaut, in press). 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 1.

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. We are currently investigating whether minimizing the relative frequency of the sticky pairs leads to a network with more stable behavior.

Following Rohde and Plaut (1999), 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 at least 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 is one class of hybrid Connectionist-Symbolic models which may, with some modifications, naturally predict Ungrammatical Influence effects: it is the class consisting of the the Competititve Attachment Processor ("CAPERS") of Stevenson (1994) and the Dynamical Unification-Space parser of Vosse and Kempen (1999) (See also Kempen & Vosse, 1989). 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 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 additonal 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 will 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 may be just what is needed to permit a parallel processing solution to the tractability problem. Thus, the signifcance of the Ungrammatical Influences hypothesis is that it pushes us in the direction of seeking such a solution.

We turn, now, to empirical investigation of the Ungrammatical Influences 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 patterns that constrained the next word to be a verb (e.g., *Some people cannot...*, *We decided to...*, *a proposal to...*, *need a truck to...*). This next word (labeled "Word 0") 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, Word 0 formed a compound with the word after it (labeled "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.

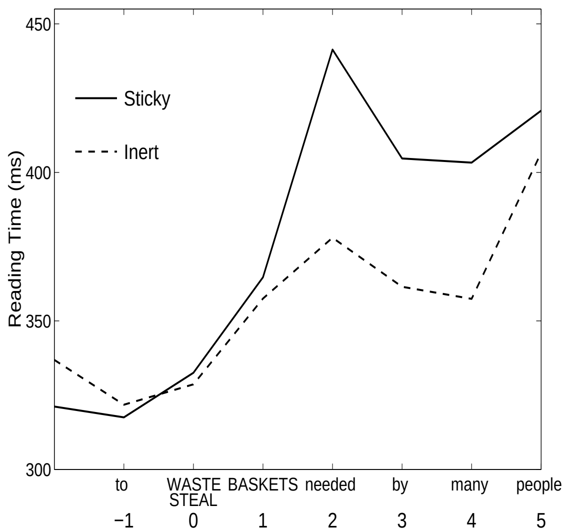
The sentences were presented using the moving-window self-paced-reading method of Just, Carpenter, and Wooley (1982): Readers ininitally see a computer screen with a sentence on it, but all the letters are replaced by dashes. They press the space bar and the first word replaces its dashes. They press again, the first word reverts to dashes, and the second word appears. When the end of the sentence is reached, the participant answers a comprehension question about the stentence. We measured the intervals between space-bar 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.

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

We computed the base 10 logarithms of the raw reading times and performed a linear regression with characters-per-word as independent variable and sub-

Figure 2: Graph of mean reading time versus position for Experiment 1.



jects as random factor. The analyses we report below were performed on the standardized residuals from this regression analysis.

Figure 2 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 (in the word-by-word, self-paced reading sense) exist.

But there is an alternative explanation of the outcome of Experiment 1 which should be considered. An early indication of the existence of Ungrammatical Influences came from a priming experiment 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 constrainiing 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 brough 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 (Mohanan, 1986). Perhaps the parser correctly chooses to treat these sequences as Noun-Verb collocations, but the lexicon is independently and continuously monitoring the speech stream for familiar units. When the lexicon notices “fire truck”, it creates some interference for the parser.

Thus Experiment 1 does not decisively demonstrate the existence of Ungrammatical Influences. The next experiment is designed to check 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 (5) 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.

Subjects were recruited from classes and through advertisement on the campus of the University of Connecticut. All were native speakers of English. Subjects received either money or course credit for their participation. The paid subjects received $10 per hour. The for-credit subjects received 2 credits per hour. The experiment lasted for about 30 minutes.

Materials.

Eighteen experimental items were created. These are listed in Appendix 2. Each item involved four conditions as in (6):

(6) The manager watched the waiter...

		0	1	2	3	4	5
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) (6a and c) or unreduced (UR) (6b and d); its past participle verb was either homophonous and homographic (H) with the corresponding past tense form (6a and b) or distinct from it (NH) (6b and d). 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., Frazier and Rayner, 1982; Ferreira and Clifton, 1986; Trueswell, Tanenhaus, and Garnsey, 1994; Trueswell, 1996).

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

There is strong evidence that when it is semantically sensible to interpret the verb following the subject noun as 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 ("rented"). If readers were to compute such a reading, then, this would be a case of an ungrammatical influence.

Thus, on the hypothesis that the sequence "the waiter served pea soup" creates a distracting ungrammatical influence, we expect that subjects will be slower at reading the Reduced/Homophonous condition (a) starting at or shortly after the word "served" than they will be at reading the Reduced/Nonhomophonous condition (c) starting at the word "shown". However, in a direct comparison between cases (a)and (c), it would be hard to rule out the possibility that a contrast between these two cases was due to independent semantic or lexical frequency differences between two verbs. Therefore, we included conditions (b) (Unreduced/Homophonous) and (d) (Unreduced/Nonhomophonous), which have identical verbs, but which are disambiguated periphrastically. The prediction of the Ungrammatical Influences hypothesis, then, is that there will be a significant interaction between the factors Reduction and Homophony, with reduction speeding reading time more in the homophonous case than the contrastive case.

We are looking for an effect of Reduction for the homophonous case. If this effect obtains and the reduced cases are read faster than the unreduced 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 an speed-enhancing 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.

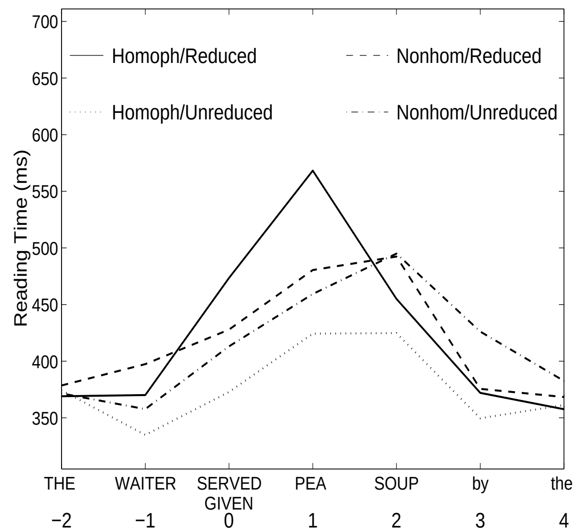
Procedure

The procedure was the same as for Experiment 1.

Results. All subjects scored better than 90% 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 a two factors: Homophony and Reduction. There was a main effect of Reduction in the region defined by Words 0 to 3. Homophony did not produce a significant main effect anywhere. In the subject analysis and the item analysis, there was a signifcant interaction between Homophony and Reduction over the region defined by words 0 to 3. There interaction was also significant at Word 0, Word 1, and Word 2. A graph of scaled reading times for Experiment 2 is shown in 3. A graph of the interaction is shown in Figure 4. As the figure indicates, reduction slowed reading times in both the homophonous and the nonhomophonous conditions, but the slowing is significantly greater in the homophonous case.

Figure 3: Reading times in the four conditions of Experiment 2.

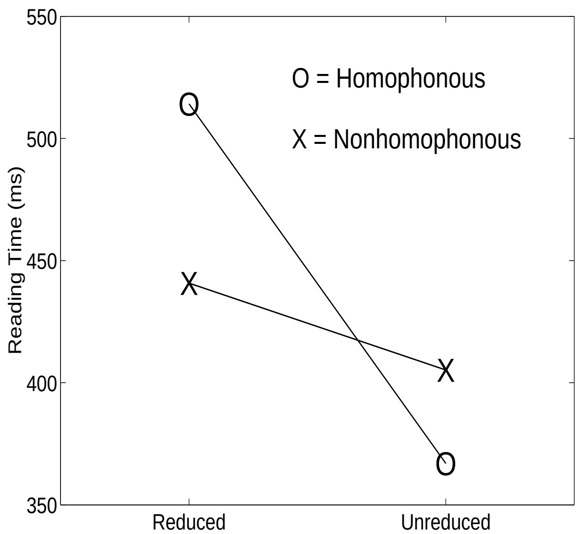


Discussion

The existence of the interaction, with the reduction slowing the homophonous case, more than the distinct 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 in the following sense. In experiment 2, the local structural ambiguity has the form NP V NP (e.g. the waiter served pea soup). The effect begins on Word 0 ("served") and disappears by about Word 3 (the word "by" in most of the examples). In Experiment 1, the local structural ambiguity has the form N N (e.g. "fire truck"). In this case, the first Noun is Word 0, and the local ambiguity is over by Word 1, but the effect does not become significant on an individual word until Word 2 (and it remains significant over words 3 and 4). We speculate that this difference in the timing of manifestation of Ungrammatical Influences across the two experiments may stem from the fairly unusual syntax of the grammaticality correct interpretation of the Experiment 2 sentences. Reduced relative structures are rare in the first place. Reduced relative structures with ditransitive verbs are especially unusual, so readers may be working hard to interpret the sen-

Figure 4: Interaction between Homophony and Reduction in Experiment 2.



tences in the first place, and an additional distraction from an Ungrammatical Influence may thus easily disrups processing. By contrast, the syntactic structures of Experiment 1 are very common modal+Infinitive or "to"+Infinitive collocatios, so readers may not tend to detect the distracting influence until it has had time to "sink in" more.

Conclusion

We have focused on the hypothesized phenomenon of Ungrammatical Influences: the syntactic parser is expected to be influenced by local, phrasal coherence that is incompatible with the structure of preceding syntactic material. Two experiments supported the existence of Ungrammatical Influences in parsing. We noted that such effects push the theory of parsing strongly in the direction of dynamical, self-organizing models.

Although the present experiments suggest treating Ungrammatical Influences as a kind of interference effect (consistent with the general class of Limited Resource models of parsing), in other work, we have found that Ungrammatical Influences do not always get in the way of parsing. Galanutucci, 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.*), parsing was facilitated, even though grammatically, the binding is disallowed. These results indicate that we need to probe more deeply into the problem of incoherent sturcture 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 Univeristy of Connecticut Linguistics/Psychology lunch talk group.

Elman, J. L. (1990). Finding structure in time. Cognitive Science, 14, 179–211.

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

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

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.

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

Kempen, G. and Vosse, T. (1989). Incremental syntactic tree formation in human sentence processing: a cognitive architecture based on activation decay and simulated annealing. Connection Science, 1:273-290.

Rohde, D. L. T., & Plaut, D. C. (1999). Simple recurrent networks can distingish non-occurring from ungrammatical sentences given appropriate task-specific structure: reply to Marcus. Cognition 73(3), 297-300.

Rumelhart, D. E., Durbin, R., Golden, R., and Chauvin, Y. (1995). Backpropagation: The basic theory. In D.E. Rumelhart and 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 (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. and Richardson, D. (1999). 'Ungrammatical influences in sentence processing'. Poster presented at the 12th Annual CUNY Sentence Processing Conference, New York, NY.

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