Ungrammatical influences in sentence processing

Whitney Tabor
University of Connecticut

tabor@uconnvm.uconn.edu
http://www.sp.uconn.edu/~ps300vc/tabor.html

Daniel Richardson
Cornell University

dcr18@cornell.edu

March 18, 1999
CUNY Graduate Center, New York

Introduction

Evidence that many kinds of linguistic information (syntactic, lexical, discourse, etc.) can bias parse choices has led to the development of interactionist models of parsing (see Tanenhaus and Trueswell, 1995, for review). Among the interactionist models, one subgroup, connectionist models, make bottom-up structural organization primary, and thus make an additional, surprising prediction about the parsing of ambiguous elements: even grammatically impossible interpretations, if they find sufficient local motivation, may exert an influence on the parsing process. The importance of this claim is that it points to a much richer set of relationships between words and phrases than current syntactic theories, which focus on well-formedness, provide. We must ask: where, in the model of grammar and/or processing, should the ungrammatical structures reside?

Evidence for Ungrammatical Influences at the Lexical Level

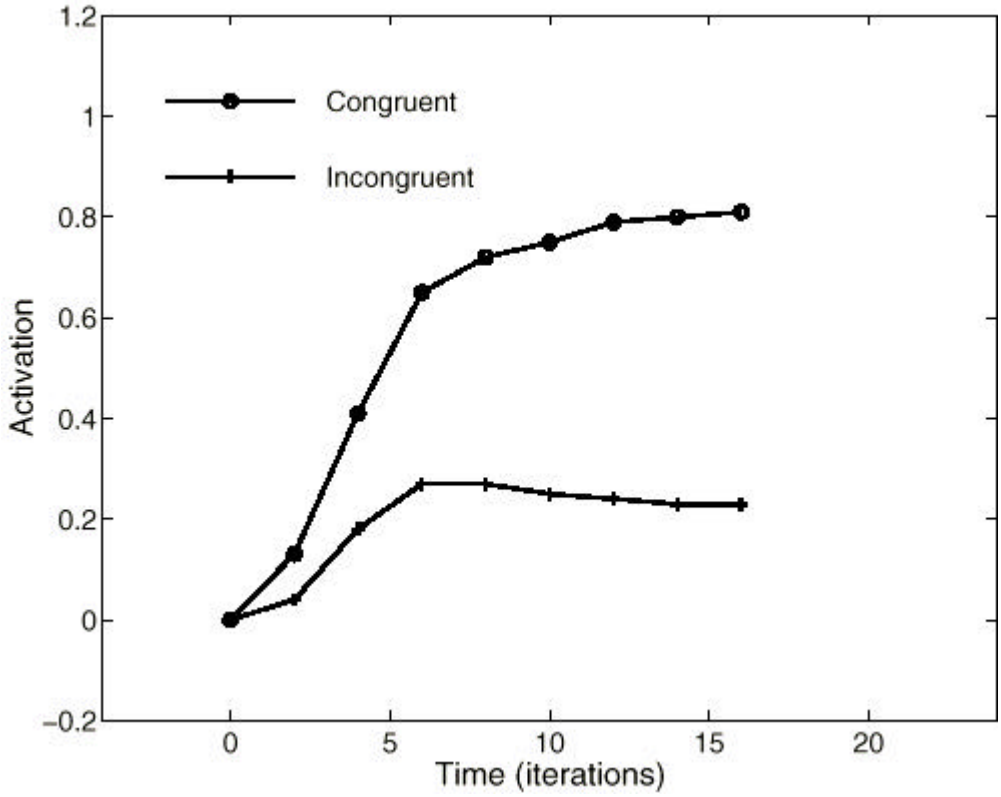
Evidence in favor of one simple variety of the "ungrammatical influences" claim has been known for some time. Tanenhaus, Leiman, and Seidenberg (1979) found that both the noun and verb interpretations of ambiguous lexical items (e.g., "rose") temporarily (for less than 200ms after presentation) support priming even in a syntactic context which rules out one interpretation (e.g., "She held the rose; They all rose") (Fig. 1). Although the data at 600 ms. are somewhat complex, there is a clear development from activation of both senses ("multiple-activation") to activation of only the syntactically relevant sense during the 0-200ms interval. The Tanenhaus et al. results parallel the results of Swinney (1979) who showed that both meanings of a semantically ambiguous word (e.g., "bug") were activated for a short period of time even in a semantically biasing context.

Fig. 1. Activation of two syntactically distinct senses of ambiguous words during sentence processing. In the figure "Cat_1-Cat_2" refers to the

Models of early multiple-activation

When early multiple-activation of ambiguous words was first detected, it was perceived as evidence for a modularity thesis (e.g., Tanenhaus, et al. 1979). Evidently, lexical information associated with a word exerts an influence on processing independently of the syntactic context in which the word appears during an initial, brief stage of processing. Interestingly, highly "interactive" models, like connectionist networks can predict multiple-activation as well. They implement a kind of "soft" modularity by assigning different units to word and context representations. Kawamoto (1993) successfully simulated early multiple activation in a dynamical connectionist network (Fig. 2).

Fig. 2. Initial activation of both (semantic) senses followed by suppression of the contextually incongruent sense in a dynamical neural network model (From Kawamoto, 1993, Fig. 9.)



Beyond the one-word level

But do such ungrammatical influences ever extend past the one-word level and the 200ms time interval? As a first step toward answering this question, we used a word-by-word, self-paced reading paradigm to compare sentences like (a) below, which has an embedded two-word compound (*waste baskets*) that is irrelevant to the only reasonable parse, with sentences like (b) which has no such distractor.

Word position: -1 0 1 2 3 4
(a) It is not wise to waste baskets needed by many people.

Word position: -1 0 1 2 3 4
(b) It is not wise to steal baskets needed by many people.

Results

Reading times at positions 2, 3, and 4 were significantly higher in the (a) cases (Fig. 3a) than the (b) cases (Fig. 3b) ($F1(1, 30) = 10.77$, $p = .003$; $F2(1, 15) = 4.79$, $p = .045$), thus supporting the hypothesis that there is interference from the ambiguous, but grammatically irrelevant, compound.

Fig. 3a. Multi-word interference effect (Subject Analysis). The figure compares reading times for sentences containing distractor compounds ("Sticky") with reading times for sentences lacking such compounds ("Inert"). $F1(1, 30) = 10.77$, $p < .005$ for word positions 2, 3, and 4 together.

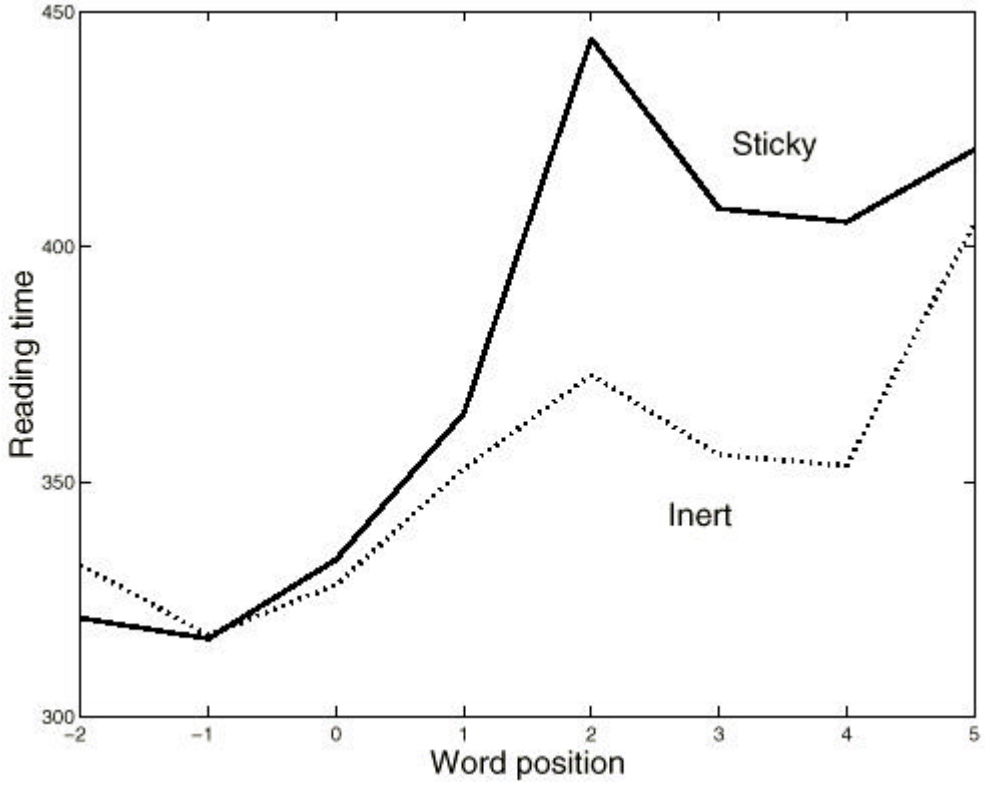
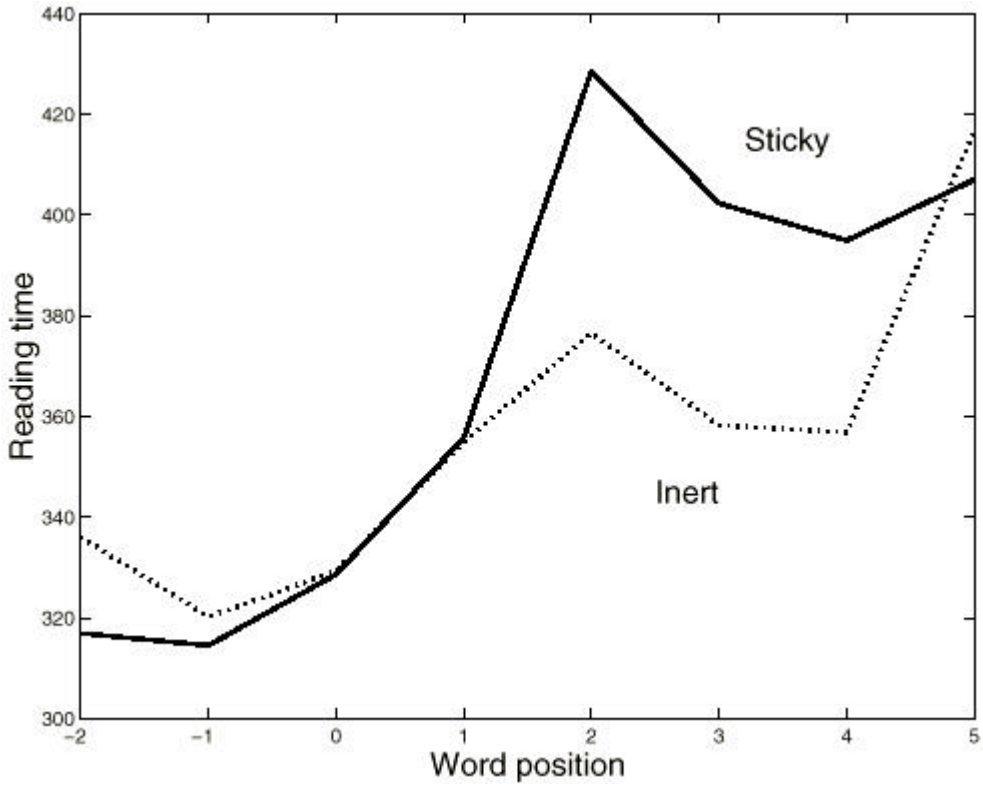


Fig. 3b. Multi-word interference effect (Item Analysis). $F2(1, 15) = 4.79$, $p < .05$, for word positions 2, 3, and 4 together.



A Mental Simulation Experiment

Consider a model which develops syntactically predictive representations in a metric space by starting with a tight random cluster and bifurcating, with information tending to diffuse across time. Symbol processing connectionist networks (e.g. Elman, 1990; 1991) are of this type.

The data are given by the grammar in Fig. 4. This grammar approximates the situation tested in the experiment in that there are two distinct syntactic contexts each of which selects one interpretation of an ambiguous two word sequence: Verb Noun_{Obj} or Noun Noun (compound). The job of the model is to predict which words are most likely to come next at any point in a sequence.

Since information tends to diffuse across time, early learning capitalizes on correlations between nearby items. Thus, there is a stage in learning when the model develops t representations based on a trigram window. Fig. 5a shows the region of the space associated with sequences of the form "the Noun..." during this phase. The figure is a two dimensional projection of a high dimensional representation in which each Noun commands a separate dimension. The model learns this clustering in order to accurately predict the second Noun. We assume, consistent with metric representation principles, that correlated items induce correlated representational changes (so "waste" and "baskets" tend to add but "waste" and "mats" tend to cancel). Fig. 5b shows the resulting projection for the region associated with "the Noun Noun...". There is an

essential distinction between a centrally located cloud of points and a peripheral circle. The cloud consists of cases where the two nouns don't stick (e.g. "waste mats"). The circle is actually a ring of clusters, each corresonding to a lexicalized compound (e.g., "waste basket"). Eventually, 4-gram conditioning is achieved, and the ring in Fig. 5b separates into two concentric rings (Fig. 5c), the outer one corresponding to the compound context (introduced by "the") and the inner one corresponding to sticky pairs in the infinitive context (introduced by "to").

Although the distinction between the inner ring and the cloud is no longer needed, it remains as a residue of the earlier state, causing displacements from the optimal (mean) representation. These displacements correspond to longer reading times in the contextually inappropriate sticky pairs under either a map from prediction error to time (Christiansen and Chater, in press) or a dynamical implementation (e.g., Tabor, Juliano, and Tanenhaus, 1997; Tabor and Tanenhaus, in press).

Fig. 4. Training grammar for the mental simulation experiment.

$S \rightarrow S_{VP}$
$S \rightarrow S_{NP}$

$S_{VP} \rightarrow$ to V N_{Obj} is unforgivable
V $\rightarrow$ waste, place, bear, cart, fuel, mail
$N_{Obj} \rightarrow$ baskets, mats, cubs, wheels, tanks, men

$S_{NP} \rightarrow$ the waste baskets are large
$S_{NP} \rightarrow$ the place mats are flat
$S_{NP} \rightarrow$ the bear cubs are round
$S_{NP} \rightarrow$ the cart wheels are shaky
$S_{NP} \rightarrow$ the fuel tanks are full
$S_{NP} \rightarrow$ the mail men are persistent

Fig. 5a. Metric space representation at the point of processing the second word (Noun or Verb) of a test sentence early in training (e.g. "the/to waste...").

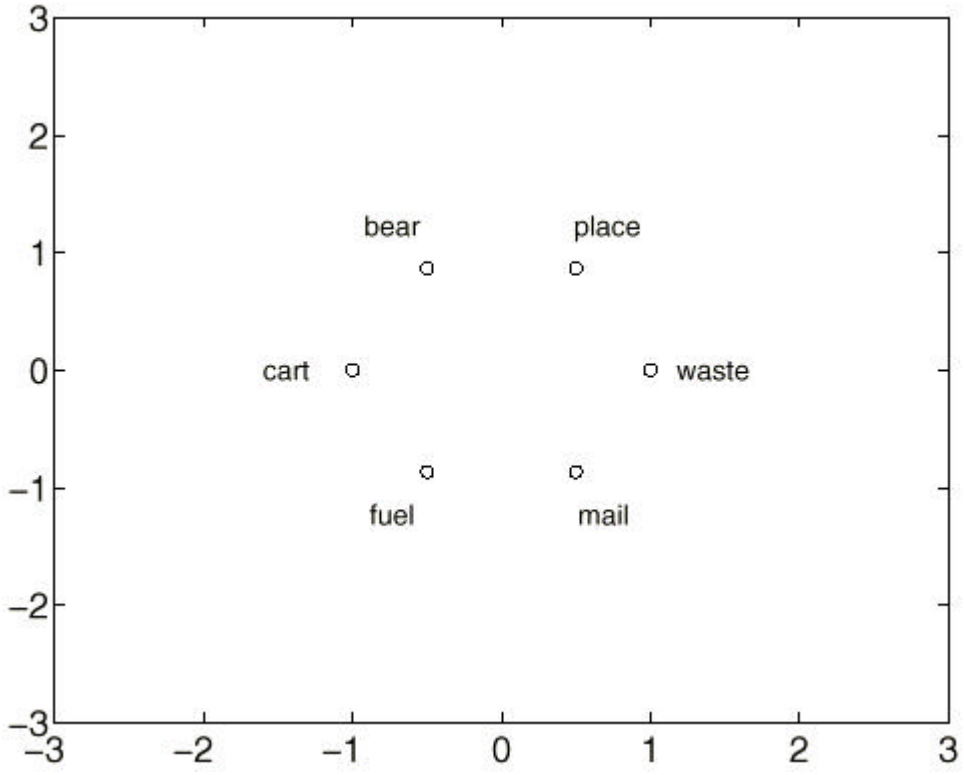


Fig. 5b. Metric space representation at the point of processing the third word (necessarily a Noun) of a test sentence somewhat later in training. (e.g. "the/to waste baskets...").

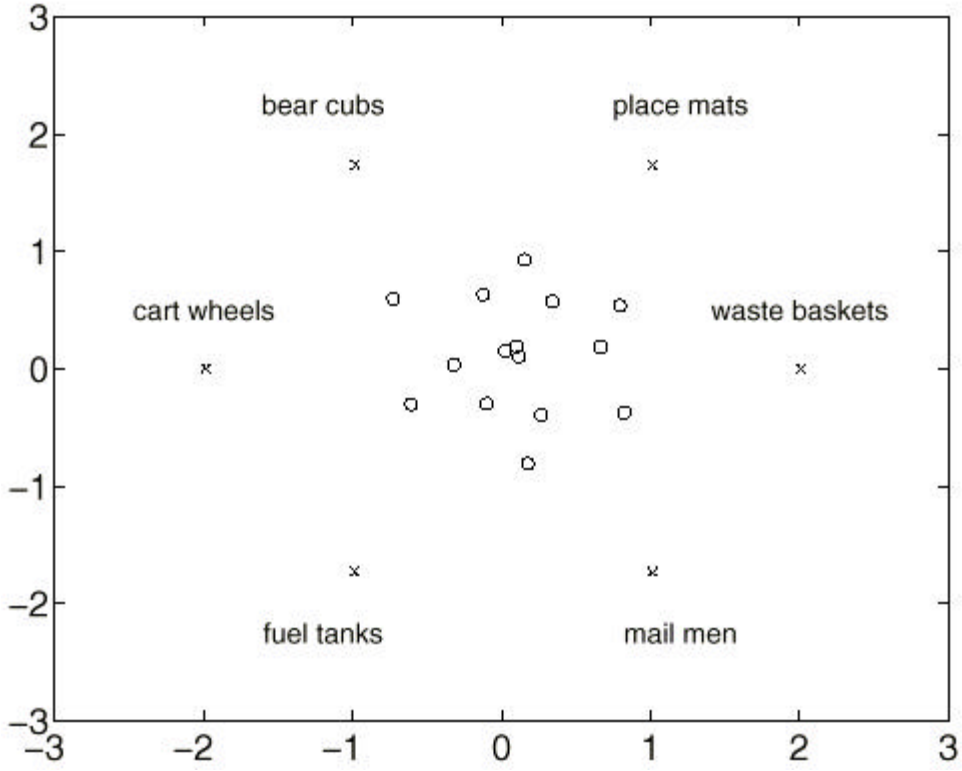
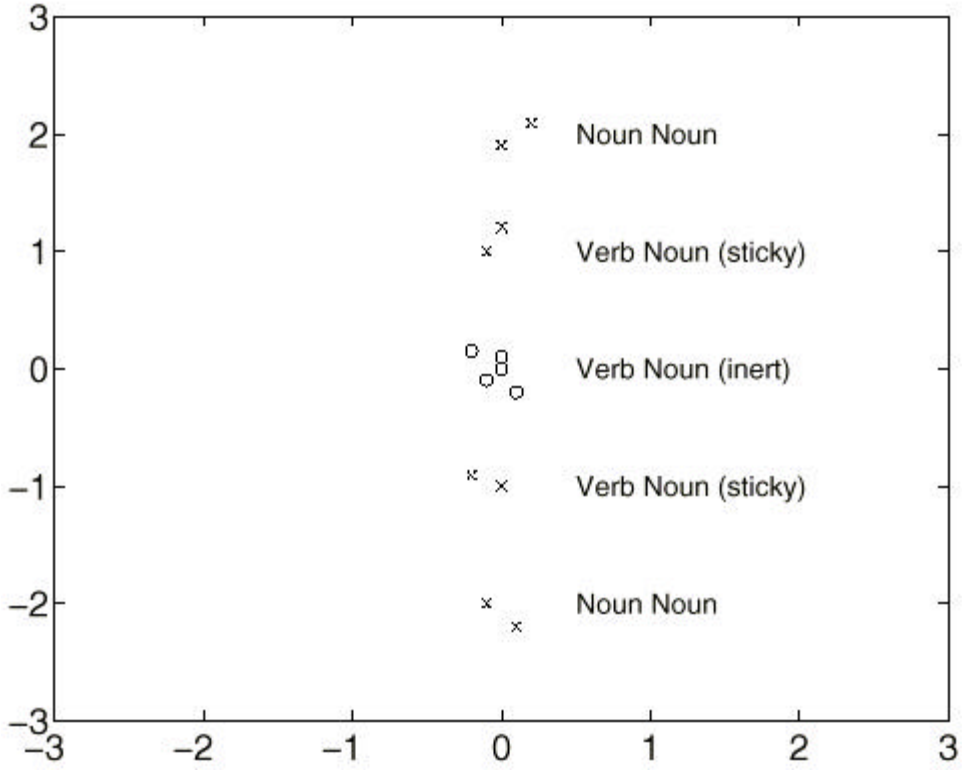


Fig. 5c. Metric space representation at the point of processing the third word (necessarily a Noun) of a test sentence when the grammar has been fully learned.



Conclusions

These results are challenging for models of parsing based on the construction of phrasal units. From a tree-building standpoint, the results of the current experiment suggest that trees are not built strictly from left to right, but that small subtrees inconsistent with the main tree may temporarily form and then disband. The chemical bonding grammar model of Kempen and Vosse (1989) looks like a promising revision of the tree-building approach from this standpoint.

The mental simulation experiment shows how a metric space model like a connectionist network predicts the ungrammatical influence effects. In essence, influence from spurious local associations stems from earlier stages of learning when local associations had predictive value.

An important empirical question is whether such effects extend to longer sequences of distractors and to distractors that are not lexical compounds.

Our results suggest that "interaction" is only the tip of an iceberg: the theory of grammar and processing must also support an appropriate "intimacy" between the well-formed structures that compose a language and a ghostly legion of quasi-structures whose members are sometimes temporarily entertained.

References

Christiansen, M.~H. and Chater, N. (in press).
Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science*.

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

Kawamoto, A.H. (1993). Nonlinear dynamics in the resolution of lexical ambiguity: A parallel
distributed processing account. *Journal of Memory and Language*, 32:474--516.

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.

Swinney, D. A. (1979). Lexical access during sentence comprehension: (re)consideration of context effects. *Journal of Verbal Learning and Verbal Behavior*, 18:645--660.

Tabor, W., Juliano, C., and Tanenhaus, M. (1997).
Parsing in a dynamical system: An attractor-based account of the
interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes*, 12(2/3):211--271.

Tabor, W. and Tanenhaus, M. (In Press).
Dynamical models of sentence processing. *Cognitive Science*.

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., editors, *Handbook of Perception and Cognition: Volume 11*, pages 217--262. Academic Press, San Diego.

Appendix: Stimuli

1. He had to run/jog way across the field to fetch the missing jacket.

2. We decided to place/lay mats on the floor before we did yoga.

3. It's not wise to waste/steal baskets needed by many people.

4. Some people cannot bear/stand cub scouts under the age of 12.

5. The government did not think it could police/control men who had spent more than ten years in the territories.

6. On Sunday, the whole family would watch/attend dog races in the country.

7. The company decided to fly/ship paper daily to its manufacturing plant in Brazil.

8. We don't think they fire/hire truck drivers without consulting the union.

9. The Filipino workers didn't like having to machine/construct parts for the American tanks.

10. In the evenings along the boulevard, the young men would sport/don jackets they had purchased from the truckers.

11. After the scandal became public, the FBI began to tail/track pipe smokers from Columbia and Venezuala.

12. The company did not think its gas tanks would fail/flunk safe storage test that they had pased many times before.

13. The town did not think it could stop/keep signs from being vandalized by arresting teenagers.

14. Under the new constitution, the policement were required to state/cite laws justifying their actions whenver they arrested somebody.

15. The bicycle builders will need a 20 foot truck to cart/haul wheels into the city for the technology fair.

16. The army was doing research on a proposal to fuel/power tanks with methane when budget cuts nixed the project.