

Helping a jury understand witness deception

David Singleton

University of Texas at Arlington

In Stephen Vincent Benet's short story "By the Waters of Babylon," the chief of a tribe comes to the place in his life in which he needs to pass on the values of the tribe to his son who will take his place leading the tribe. So he sends his son on a journey in hopes that the boy will develop realizations about life that will enable him to become a wise leader. After the boy's return, the chief sits down to discuss his son's new understandings. The son speaks mainly about discovering certain truths in life that are different from the truths that his father has believed. The father acknowledges this relative nature of truth with the memorable line, "Truth is a hard deer to hunt" (Benet 1937). I want to make this memorable line the theme of this paper, applying it to the courtroom, however, rather than to tribal succession.

In our own business-oriented society, men and women represent matters to each other routinely. For one reason or another, the matters represented become hazy or twisted or remain unfulfilled. Lawsuits result. Advocates for the two parties involved then try to arbitrate the parties' differences by seeking for the original representations made. In preparations for trials and in trials themselves, attorneys try to reconstruct what the truth of the matter may have been. But, in the courtroom, juries who will decide the outcomes for the disputing parties realize, like the chief, that "truth is a hard deer to hunt." The different parties have different understandings of the representations made, so it is hard to discover the truth.

Because of differences in understandings, our society has made it a priority to design both devices and methods to arrive at the truth of a matter. People take lie detector tests or have truth serums administered or undergo a battery of tests, interrogations, or evaluations. Linguists, too, have entered the arena and devised methods for discovering the truth of a matter. Two of these methods are the subject of this paper. Shuy's topic flow analysis and Carpenter's type-token ratio application will be examined for the value they have in determining witness deception either through lying or through making statements serving hidden purposes (*Machiavellian intent* according to Carpenter). The two methods, however, will be examined only to the extent that they illustrate a certain compatibility with each other. It is at the point of their compatibility that these methods enhance their usefulness in determining the likelihood of deception of the parties involved.

Content made available by
Georgetown University Press,
Digital Georgetown, and
the Department of Languages and Linguistics.

Shuy's topic flow analysis, from his 1986 article, evidenced the cooperation or lack of cooperation a suspect showed in taped conversations with a law enforcement agent. The topics raised by the agent were classified, isolated, and listed. The answers to those topics deemed most important, what Shuy called substantive, were examined for the suspect's positive, negative, or neutral responses. Shuy illustrated that a suspect exhibited non-cooperative language behavior in the tightly bound context of conversation and did not incriminate himself as the agent wanted listeners to the tape to deceptively believe. The particular part of the method of interest to this paper is the division of topics into the three classes, substantive, corollary, and transitional. These classes allow one to sort the conversation into areas of importance so that what an active listener merely gets an impression of, the fine-grained analysis can exploit for fine shades of meaning and precise kinds of measurement.

The second method, Carpenter's type-token ratio, operates from the researched notion that utterances generate tokens on a repetitive basis when a person is in a relaxed frame of mind. But, if a person experiences apprehension or caution, the generated tokens show greater lexical diversity. Carpenter's particular application of this principle, from his 1986 article, is to a police interrogation transcript of a murder suspect and to a defendant's courtroom testimony in a rape case. Under Carpenter's analysis, the problematic areas of testimony in the two cases coincide with the suspect's and defendant's highest type-token ratio. In fact, noting the areas of greatest lexical diversity is the portion of this method of greatest interest to this paper. In particular, the formula used for arriving at a ratio is well suited for comparisons to other methods and for evaluations in a context-free environment. Testimony of the suspect and defendant that contained utterances of at least and only fifty words were isolated to maintain statistical significance based upon English speakers' repeating a word every ten to fifteen words. The number of different words in the fifty-word utterances, called types, were then divided by the number of total words, called tokens, to provide a ratio. The higher the ratio, the greater the number of different types. However, one must provide a framework for what constitutes a high type-token ratio. Thus, all segments containing fifty-word utterances, no more and no less, are counted in order to find the mean and the standard deviation for the distribution of segments. The type-token ratios above the +1 standard deviation, and certainly above the +2 standard deviation, represent utterance areas where caution is being exercised. These areas can be investigated, then, for Machiavellian intent, self-incrimination, or well-rehearsed recountings.

One can see the merits of both methods and their usefulness in highlighting spots in a testimony used for deceptive purposes. One might also see the areas of possible overlap if the two methods were to be combined. They have a certain compatibility to them, and not often can one wed a context-free perspective with a context-bound one. In the case of these two methods, a person can ask, What

if one were to compare the fully elaborated responses to substantive topics with the type-token ratio for that response? Could a person measure relative ease in response to certain topics and apprehension or caution in response to other topics? Could the two methods combine to act as a net to snare those portions of testimony deceptive in some way?

I believe a person can capture deception in language behavior, and the technique used to capture this behavior is a mapping of the type-token ratio onto a mapping of transitional, corollary, and substantive topics. To accomplish this mapping, one must take a transcript of testimony and classify the questions and answers into their three classes. Second, one must identify all the responses of fifty words so that a mean for the speaker's type-token ratio can be derived. Following the mean calculation, one finds the +1 standard deviation threshold. Finally, one observes on what substantive topics the speaker exhibits high type-token ratios. I will illustrate this technique by using two cases supplied to me by the law firm of Cotton & Bledsoe in Midland, Texas, Rick Strange, attorney.

The first case involved an oil field worker who had been injured due to the collapse of a hydraulic arm supporting the bucket the plaintiff was standing in. The injured man sued all parties who had owned the bucket truck prior to the accident and the party who had brought the truck to its owner at the time of the injury. The transcript being evaluated was a deposition conducted by the attorney for the injured party, the plaintiff's attorney. The person being deposed was the owner of the company that located the bucket truck, Doyle, who had facilitated its sale for the company that owned the truck at the time of the injury, C&D Coating. During the questioning, the plaintiff's attorney clearly implicated Doyle's responsibility for the collapse of the hydraulic arm because Doyle knew the poor condition of the truck before showing the truck to C&D Coating. Using the technique just described, the deposition's questions and answers were classified as substantive, corollary, or transitional. The list of substantive topics is shown in Table 1.

Four substantive topics emerged. Trenchard, plaintiff's attorney, spoke of Doyle's competence, others' reliance on Doyle, the history of the bucket truck, and Doyle's representations to C&D Coating. Next, the utterances containing at least and only fifty words were identified. Thirty-five such segments existed. They were ranked and a mean was calculated. The calculations and rankings are seen in Table 2. The +1 standard deviation was figured and plotted on a graph showing the ratio points as they occurred in the chronological order of the deposition. These points portray peaks above the first standard deviation, represented in Figure 1. Next, a linear map was made to depict the topics in the order they were introduced in the deposition shown in Figure 2. Finally, a linear representation of the type-token ratios was drawn below the topic map, shown in Figure 3.

This last linear map shows that Doyle performed at a level of relative ease during the topic of how he had represented the truck to C&D Coating. Type-

Table 1. Roger Shuy's substantive topic-flow analysis

Doyle	Topic No.	Trenchard
Responds fully	9	Others' reliance on Doyle
Responds fully	11	Others' reliance on Doyle
Responds fully	26	Question representations made
Responds positively	29	Others' reliance on Doyle
Responds fully	30	Question representations made
Responds fully	31	Full history disclosure
Responds fully	34	Testing Doyle's competence
Responds negatively (follows with full response)	42	Testing Doyle's competence
Responds positively (follows with full response)	50	Testing Doyle's competence
Responds negatively (twice) (follows with full response)	56	Full history disclosure
Responds fully (to the contrary)	58	Questions representations made
Responds fully (to the contrary)	59	Questions representations made
Responds positively	60	Questions representations made
Responds fully (follows to the contrary)	63	Testing Doyle's competence

token ratios for this line of questioning range from .40 to .64. The type-token ratios for the idea that Doyle should have looked into the history of the rig are raised twice. The first time it is raised, Doyle seems at ease (.64), but the second time he feels the weight of what an incriminating position the subject places him in, so his answers register .70 and .76. Another topic, Doyle's com-

Table 2. Ronald Carpenter's type-token ratio

Rank order of segments by TTR order					
TTR	Segment No.	Rank	TTR	Segment No.	Rank
.82	5	1	.60	9	10
.80	4	2	.60	13	10
.78	31	3	.60	16	10
.76	3	4	.60	28	10
.76	17	4	.56	8	11
.76	26	4	.56	29	11
.72	15	5	.56	30	11
.70	25	6	.54	19	12
.68	18	7	.54	35	12
.64	10	8	.52	27	13
.64	11	8	.50	2	14
.64	14	8	.46	1	15
.64	21	8	.44	24	16
.64	23	8	.42	12	17
.64	32	8	.42	34	17
.64	33	8	.40	7	18
.62	6	9	.40	20	18
.62	22	9			

Notes: 35 segments of 50 words.

Mean TTR (Type-Token Ratio) = .61 (.6065).

+ 1 SD (Standard Deviation) = .11 (.1143) or [.72 in the distribution].

+2 SD = .22 or [.83 in the distribution].

Content made available by
Georgetown University Press,
Digital Georgetown, and
the Department of Languages and Linguistics.

Figure 1. Ratio peaks: A defendant's comfort and discomfort zones

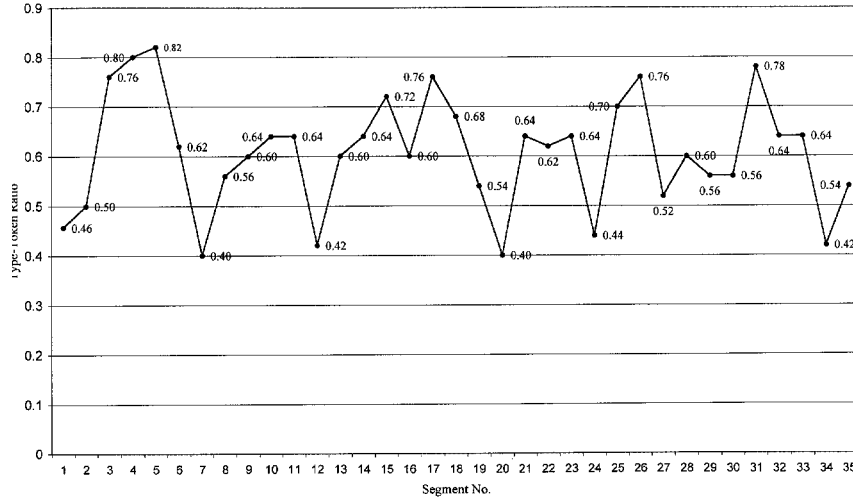
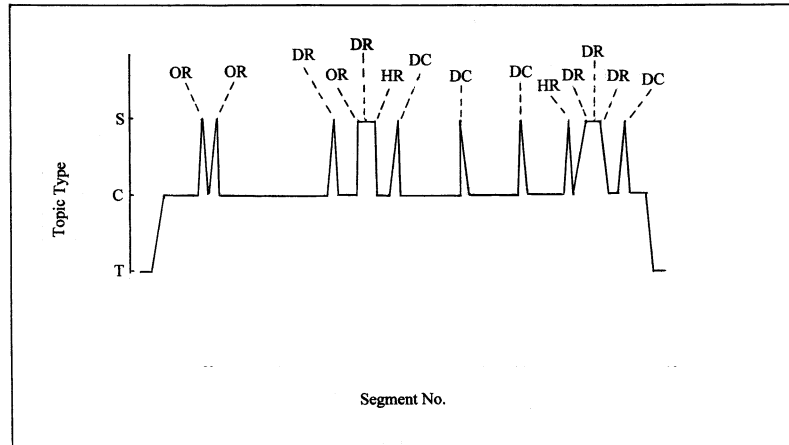
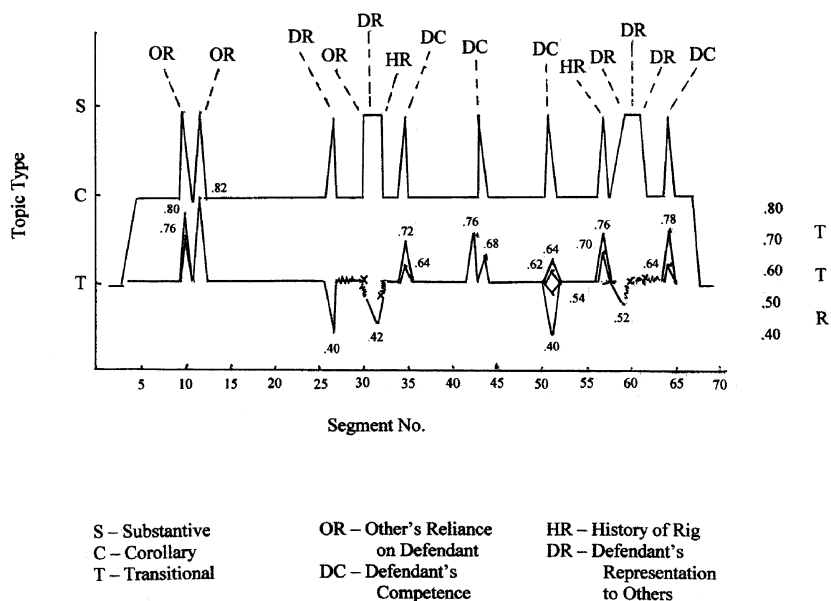


Figure 2. Substantive question categories: Mapping a questioner's categories



S – Substantive
 C – Corollary
 T – Transitional
 OR – Other's Reliance on Defendant
 DR – Defendant's Representation to Others
 HR – History of Rig
 DC – Defendant's Competence

Content made available by
 Georgetown University Press,
 Digital Georgetown, and
 the Department of Languages and Linguistics.

Figure 3. Substantive question categories: Mapping a questioner's categories

petence, comes into question on four occasions. Three out of four of these responses are at or above the first standard deviation. The first time the topic is raised, he has two responses (.64, .72). Apparently, he was unsure of Trenchard's strategy. The second mention of his competence makes Doyle aware that what he says might be held against him. The third instance of questioning his competence receives a fully elaborated response, but Doyle seems to feel comfortable with either his perceived position or the direction the questioning has taken. However, during the final occasion for the topic, he registers his highest type-token ratio for the topic (.78). Doyle's caution is evident because he clearly sees the questioner's intentions for bringing up his competence so frequently and for suggesting that he is not competent to evaluate the condition of oil field equipment, which would make him liable for the accident. And last, Doyle's highest type-token ratios come at the beginning of the deposition and border on two standard deviations above the mean (.80, .82). The topic of others' reliance on him is raised three times. The third occurrence cannot be evaluated since his answer did not reach fifty words. Thus, an X is placed on the mean line. Doyle's responses to the first two instances of the topic cause great lexical diversity. He knows that if the plaintiff can prove his client's reliance on him, then his case is lost.

As the case turned out, Doyle was not convicted. The defendant's attorney adequately argued Doyle's innocence. However, the plaintiff's attorney chose his main argument to be the topic of Doyle's representations to C&D Coating. Little did he know that that was the area of most relative ease for Doyle, and that Doyle did not see himself incriminated on that topic. If the plaintiff's attorney had analyzed Doyle's deposition, as this paper has outlined, prior to the trial, then he would have known that the two issues to probe for incrimination would have been the substantive topics of competence and C&D Coating's reliance on him. The jury instead believed Doyle, maybe because of his relative ease in handling the questioning, and refused to find against him.

The second case selected on which to test this paper's proposed analysis is a fraud case. Two men met and agreed to start a corporation to provide a specialized trucking service to oil operators. The business, known as Carbonics, became successful. One man claimed about a year later that his partner was not living up to his side of the original agreement. So, his wife and son started a corporation, CO2 Services, which in about six months' time had the equipment and contracts of Carbonics. The two partners ended up pointing fingers at each other and crying fraud. The transcript considered in this paper is the wife's courtroom testimony as she explains the reasoning behind establishing a competing corporation, CO2 Services, to her husband's company, Carbonics. According to the proposed technique, four topics emerge. These topics are seen in Table 3. They are the terms of the original agreement, the Whiteface Project, the establishment of CO2 Services, and the contracts that became CO2 Services' contracts. The transcript contained fifty-one segments of fifty words with a mean type-token ratio of .63. These segments and their ratios are shown in Table 4. One standard deviation above the mean figures to be .73. The graph in Figure 4 shows that nine points peak above the first standard deviation. Figure 5 illustrates the linear mapping of the segments in order of the questioning according to the topics raised and their corresponding type-token ratios. One can see from the mapping of the type-token ratio segments the topics that reflect moments of relative ease and the moments of apprehension. X's are much more numerous on this graph because the judge began limiting Ms. Smith's responses. X's represent an utterance under fifty words. Arbitrarily, they are placed above and below the mean line alternately to simulate the normal wave-like pattern as seen in Figure 4.

Of the four categories, the Whiteface Project is the only one not to register any kind of type-token measurement. In a case such as this, one reverts to Shuy's topic flow analysis that traces the cooperation between questioner and respondent (Table 3). If one looks there, one sees that segment 94 contains a cooperative full response and segment 139 reflects a negative response. A questioner might want to pursue the negative response further. Responses to the second topic

Table 3. Roger Shuy's substantive topic flow analysis

Smith	Topic No.	Browning
Responds fully	26	Carbonics Original Agreement
Responds fully	27	Carbonics Original Agreement
Responds positively	29	Carbonics Original Agreement
Responds fully	34	Carbonics Original Agreement
Responds negatively	36	Carbonics Original Agreement
Responds positively	42	Carbonics Original Agreement
Responds positively	47	Carbonics Original Agreement
Responds fully	53	Carbonics Original Agreement
Responds fully	65	Carbonics Original Agreement
Responds negatively	68	Carbonics Original Agreement
Responds positively	70	Carbonics Original Agreement
Responds fully	79	Carbonics Original Agreement
Responds fully	81	Formation of CO2 Services
Responds fully	89	Formation of CO2 Services
Responds fully	94	Whiteface Project
Responds positively	106	Contract Distinction
Responds fully	115	Contract Distinction
Responds fully	117	Contract Distinction
Responds fully	120	Contract Distinction
Responds negatively	124	Contract Distinction

continued next page

Table 3. Roger Shuy's substantive topic flow analysis (*continued*)

Smith	Topic No.	Browning
Responds positively	125	Contract Distinction
Responds positively	126	Contract Distinction
Responds positively	129	Formation of CO2 Services
Responds negatively	139	Whiteface Project
Responds fully	146	Carbonics Original Agreement
Responds negatively	150	Carbonics Original Agreement
Responds negatively	152	Formation of CO2 Services

Table 4. Ronald Carpenter's type-token ratio

Rank order of segments by TTR order					
TTR	Segment No.	Rank	TTR	Segment No.	Rank
.80	16	1	.62	13	10
.80	22	1	.62	18	10
.78	30	2	.62	24	10
.78	37	2	.62	28	10
.78	51	2	.62	33	10
.76	43	3	.60	4	11
.74	5	4	.60	19	11
.74	20	4	.60	44	11
.74	42	4	.58	8	12

continued next page

Table 4. Ronald Carpenter's type-token ratio (*continued*)

TTR	Segment No.	Rank	TTR	Segment No.	Rank
.72	15	5	.58	12	12
.72	47	5	.56	10	13
.70	7	6	.56	26	13
.70	25	6	.56	38	13
.70	29	6	.54	11	14
.70	48	6	.54	17	14
.70	49	6	.54	31	14
.68	1	7	.54	36	14
.68	2	7	.54	41	14
.68	9	7	.54	45	14
.68	39	7	.52	32	15
.66	23	8	.50	6	16
.64	27	9	.50	21	16
.64	50	9	.50	34	16
			.50	46	16
			.48	14	17
			.48	35	17
			.44	3	18
			.44	40	18

Notes:

51 segments of 50 words.

Mean TTR (Type-Token Ratio) = .63.

+1 SD (Standard Deviation) = .10 or [.73 in the distribution].

+2 SD = .20 or [.83 in the distribution].

Content made available by
Georgetown University Press,
Digital Georgetown, and
the Department of Languages and Linguistics.

Figure 4. Ratio peaks: A defendant's comfort and discomfort zones

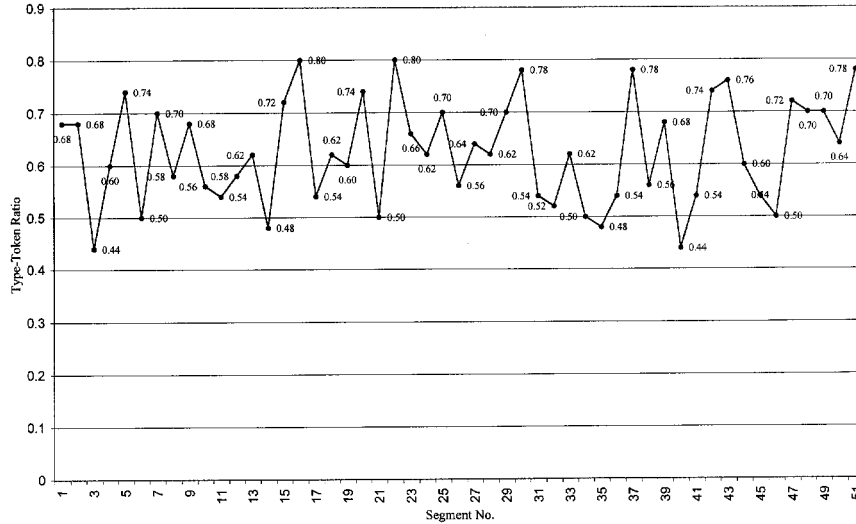
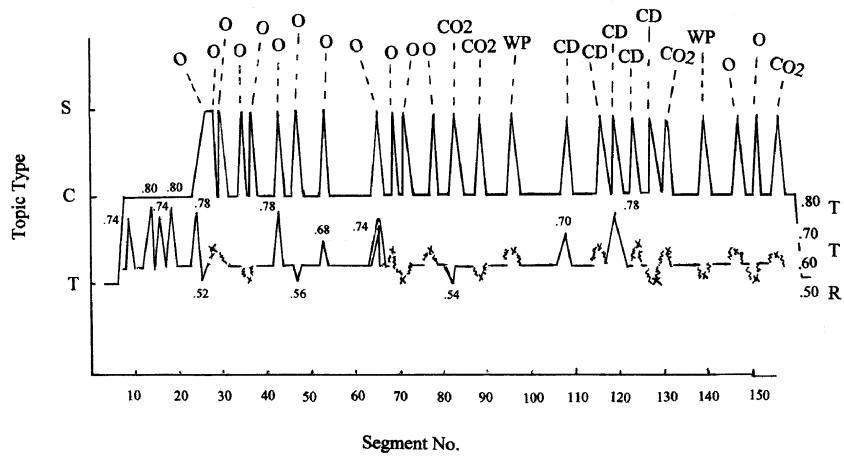


Figure 5. Substantive question categories: Mapping a questioner's categories



- S – Substantive
- C – Corollary
- T – Transitional
- O – Original Carbonics Agreement
- CO2 – Formation of CO2 Services
- WP – Whiteface Project
- CD – Contract Distinction

Content made available by
 Georgetown University Press,
 Digital Georgetown, and
 the Department of Languages and Linguistics.

of Formation of CO2 Services are barely improved from the Whiteface Project topic. Only one of four responses is measurable. However, the one measured response received a ratio of .54. If it is representative of the other three responses, Ms. Smith does not feel incriminated on the topic or does not find it necessary to give a careful or rehearsed answer. Segments 81 and 89 of Table 3 reflect full responses that tend to give the impression that the person is not hiding anything on this second subject. The line of questioning, to pursue the topic of Contract Distinction, yields two responses: one that approaches the +1 standard deviation and one that exceeds it. This topic Ms. Smith knows very well to be cautious about because not only was she privy to her husband's arrangements with his partner, but she held a corporate office in both Carbonics and CO2 Services. The type-token ratios increase to mirror the incrimination. The final topic of Ms. Smith's understanding of the original agreement between her husband and his partner yields two of the highest type-token ratios. Ms. Smith knows that if her testimony is not worded perfectly, then she crosses the line from not knowing too much about the Carbonics arrangement to knowing so much that she deliberately established the new CO2 Services to steer business away from Carbonics with its feuding partners. The two .78 ratios indicate the fine line that she is carefully walking.

One cannot help but notice that the two highest ratio peaks, in addition to two other points above the first standard deviation, surface before Trenchard begins his line of questioning on the substantive topics. He questions the defendant on topics that lead to his substantive topics. These are what Shuy has labeled corollary topics. In particular, the attorney is asking about the defendant's work and school experience, part of which included bookkeeping and economics courses. The defendant knows that by revealing the nature of her experience to the jury, they would expect her to know about transfer of assets and conflict of interest. In fact, the jury must have regarded this experience as incriminating; near the end of her testimony Ms. Smith states that she signed a document as a clerical secretary not a corporate secretary, even though the corporate seal and notary certification accompany her signature as corporate secretary. Comments from two jury members following the trial indicated that some of the jury members thought the defendant was fully aware of her conflict of interest. Ms. Smith unsuccessfully defended her position.

While truth may be a hard deer to hunt, neither is it impossible to track down. With the use of a high-powered scope and a steady trigger finger, the deer will fall in its tracks. Linguists have created the scope through which a thoughtful eye may focus on truth and deception. The use of these methods would steady a case should the attorney decide to use them. For the moment, deer hunting is seasonal.

REFERENCES

- Benet, S. V. 1937. "By the waters of Babylon." In *Selected works of Stephen Vincent Benet*. New York: Holt Rinehart and Winston.
- CO2 Services v. Carbonics Inc. et al.* 1999. Cause Number A-92,842. 70th District Court, Ector County, Texas.
- Carpenter, R. 1986. "The statistical profile of language behavior with Machiavellian intent or while experiencing caution and avoiding self incrimination." *Annals of New York Academy of Sciences*. 5-17.
- Monserat Cervantes and wife v. Dennis Doyle D/B/A Maintenance Rig Sales & Service.* 1996. Cause Number 13,508. 109th District Court, Andrews County, Texas.
- Shuy, R. 1986. "Evidence of cooperation in conversation." *Annals of New York Academy of Sciences*. 85-105.