Verification and Implementation of Language-Based Deception Indicators in Civil and Criminal Narratives

Conference Paper · January 2008

DOI: 10.3115/1599081.1599087 · Source: DBLP

3 authors:

Joan Bachenko
Linguistech LLC

Eileen Fitzpatrick
Montclair State University

Michael Schonwetter

Verification and Implementation of Language-Based Deception Indicators in Civil and Criminal Narratives

Joan Bachenko
Deception Discovery Technologies
Oxford, NJ 07863
jbachenko@comcast.net

Eileen Fitzpatrick
Montclair State University
Montclair, NJ 07043
fitzpatricke@mail.montclair.edu

Michael Schonwetter
Deception Discovery Technologies
Minneapolis, MN 55416
mschonwetter@synchronvideo.com

Abstract

Our goal is to use natural language processing to identify deceptive and non-deceptive passages in transcribed narratives. We begin by motivating an analysis of language-based deception that relies on specific linguistic indicators to discover deceptive statements. The indicator tags are assigned to a document using a mix of automated and manual methods. Once the tags are assigned, an interpreter automatically discriminates between deceptive and truthful statements based on tag densities. The texts used in our study come entirely from "real world" sources—criminal statements, police interrogations and legal testimony. The corpus was hand-tagged for the truth value of all propositions that could be externally verified as true or false. Classification and Regression Tree techniques suggest that the approach is feasible, with the model able to identify 74.9% of the T/F propositions correctly. Implementation of an automatic tagger with a large subset of tags performed well on test data, producing an average score of 68.6% recall and 85.3% precision when compared to the performance of human taggers on the same subset.

1. Introduction

The ability to detect deceptive statements in text and speech has broad applications in law enforcement and intelligence gathering. The scientific study of deception in language dates at least from Undeutsch (1954, 1989), who hypothesized that it is "not the veracity of the reporting person but the truthfulness of the statement that matters and there are certain relatively exact, definable, descriptive criteria that form a key tool for the determination of the truthfulness of statements". Reviews by Shuy (1998), Vrij (2000), and DePaulo et al. (2003) indicate that many types of deception can be identified because the liar's verbal and non-verbal behavior varies considerably from that of the truth teller's. Even so, the literature reports that human lie detectors rarely perform at a level above chance. Vrij (2000) gives a summary of 39 studies of human ability to detect lies. The majority of the studies report accuracy rates between 45-60%, with the mean accuracy rate at 56.6%.

The goal of our research is to develop and implement a system for automatically identifying deceptive and truthful statements in narratives and transcribed interviews. We focus exclusively on verbal cues to deception for this initial experiment, ignoring at present potential prosodic cues (but see Hirschberg et al.).

Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008), pages 41–48
Manchester, August 2008

In this paper, we describe a language-based analysis of deception that we have constructed and tested using “real world” sources—criminal narratives, police interrogations and legal testimony. Our analysis comprises two components: a set of deception indicators that are used for tagging a document and an interpreter that associates tag clusters with a deception likelihood. We tested the analysis by identifying propositions in the corpus that could be verified as true or false and then comparing the predictions of our model against this corpus of ground truth. Our analysis acheived an accuracy rate of 74.9%. In the remainder of this paper, we will present the analysis and a detailed description of our test results. Implementation of the analysis will also be discussed.

2. Studying Deception

The literature on deception comes primarily from experimental psychology where much of the concentration is on lies in social life and much of the experimentation is done in laboratory settings where subjects are prompted to lie[1]. These studies lack the element of deception under stress. Because of the difficulties of collecting and corroborating testimony in legal settings, analysis of so-called ‘high stakes’ data is harder to come by. To our knowledge, only two studies (Smith, 2001; Adams, 2002) correlate linguistic cues with deception using high stakes data. For our data we have relied exclusively on police department transcripts and high profile cases where the ground truth facts of the case can be established.

Previous studies correlating linguistic features with deceptive behavior (Smith, 2001; Adams, 2002; Newman et al. 2003, and studies cited in DePaulo et al. 2003) have classified narrators as truth-tellers or liars according to the presence, number and distribution of deception indicators in their narratives. Newman, et al. (2003), for example, proposes an analysis based on word likelihoods for semantically defined items such as action verbs, negative emotion words and pronouns. Narratives for their study were generated in the laboratory by student subjects. The goals of the project were to determine how well their word likelihood analysis classified the presumed author of each narrative as a liar or truth-teller and to compare their system's performance to that of human subjects. The analysis correctly achieved an overall distinction between liars and truth tellers 61% of the time.

Our research on deception detection differs from most previous work in two important ways. First, we analyze naturally occurring data, i.e. actual civil and criminal narratives instead of laboratory generated data. This gives us access to productions that cannot be replicated in laboratory experiments for ethical reasons. Second, we focus on the classification of specific statements within a narrative rather than characterizing an entire narrative or speaker as truthful or deceptive. We assume that narrators are neither always truthful nor always deceptive. Rather, every narrative consists of declarations, or assertions of fact, that retain a constant value of truth or falsehood. In this respect, we are close to Undeutsch’s hypothesis in that we are not testing the veracity of the narrator but the truthfulness of the narrator’s statements.

The purpose of our analysis is to assist human evaluators (e.g. legal professionals, intelligence analysts, employment interviewers) in assessing a text’s contents. Hence the questions that we must answer are whether it is possible to classify specific declarations as true or deceptive using only linguistic cues and, if so, then how successfully an automated system can perform the task. Our research makes no claim as to the cause of a speaker’s behavior, e.g. whether deception cues emerge as a function of emotional stress or excessive cognitive load.

3. Linguistic Markers of Deception

The literature on verbal cues to deception indicates that fabricated narrative may differ from truthful narrative at all levels from global discourse to individual word choice. Features of narrative structure and length, text coherence, factual and sensory detail, filled pauses, syntactic structure choice, verbal immediacy, negative expressions, tentative constructions, referential expressions, and particular phrasings have all been shown to differentiate truthful from deceptive statements in text (Adams, 2002; DePaulo et al., 2003; Miller and Stiff, 1993).

In the area of forensic psychology, Statement Validity Assessment is the most commonly used technique for measuring the veracity of verbal statements. SVA examines a transcribed interview for 19 criteria such as quantity of detail, embedding of the narrative in context, descriptions of interactions and reproduction of conversations (Steller & Köhnken, 1989). Tests of SVA

[1] We define deception as a deliberate attempt to mislead. We use the terms *lying* and *deceiving* interchangeably.

show that users are able to detect deception above the level of chance -- the level at which the lay person functions in identifying deception – with some criteria performing considerably better (Vrij, 2000). An SVA analysis is admissible as court evidence in Germany, the Netherlands, and Sweden.

In the criminal justice arena, another technique, Statement Analysis, or Scientific Content Analysis (SCAN), (Sapir, 1987) examines open-ended written accounts in which the writers choose where to begin and what to include in the statements. According to Sapir (1995) "when people are given the choice to give their own explanation in their own words, they would choose to be truthful . . . . it is very difficult to lie with commitment."

SCAN "claims to be able to detect instances of potential deception within the language behaviour of an individual; it does not claim to identify whether the suspect is lying" (Smith, 2001). As such, its goal is the one we have adopted: to highlight areas of a text that require clarification as part of an interview strategy.

Despite SCAN's claim that it does not aim to classify a suspect as truthful or deceptive, the validations of SCAN cues to deception to date (Smith, 2001; Adams, 2002) evaluate the technique against entire statements classified as T or F. Our approach differs in that we evaluate separately portions of the statement as true or deceptive based on the density of cues in that portion.

4. Deception Analysis for an NLP System

Our analysis is produced by two passes over the input text. In the first pass the text is tagged for deception indicators using a mix of automated and manual techniques. In the second pass the text is sent to an automated interpreter that calculates tag density using moving average and word proximity measures. The output of the interpreter is a segmentation of the text into truthful and deceptive areas.

4.1 Deception Indicators

We have selected 12 linguistic indicators of deception cited in the psychological and criminal justice literature that can be formally represented and automated in an NLP system. The indicators fall into three classes.

(1) Lack of commitment to a statement or declaration. The speaker uses linguistic devices to avoid making a direct statement of fact. Five of the indicators fit into this class: (i) linguistic hedges (described below) including non-factive verbs and nominals; (ii) qualified assertions, which leave open whether an act was performed, e.g. *I needed to get my inhaler*; (iii) unexplained lapses of time, e.g. *later that day*; (iv) overzealous expressions, e.g. *I swear to God*, and (v) rationalization of an action, e.g. *I was unfamiliar with the road.*

(2) Preference for negative expressions in word choice, syntactic structure and semantics. This class comprises three indicators: (i) negative forms, either complete words such as *never* or negative morphemes as in *inconceivable*; (ii) negative emotions, e.g. *I was a nervous wreck*; (iii) memory loss, e.g. *I forget.*

(3) Inconsistencies with respect to verb and noun forms. Four of the indicators make up this class: (i) verb tense changes (described below); (ii) thematic role changes, e.g. changing the thematic role of a NP from agent in one sentence to patient in another; (iii) noun phrase changes, where different NP forms are used for the same referent or to change the focus of a narrative; (iv) pronoun changes (described below) which are similar to noun phrase changes

To clarify our exposition, three of the indicators are described in more detail below. It is important to note with respect to these indicators of deception that deceptive passages vary considerably in the types and mix of indicators used, and the particular words used within an indicator type vary depending on factors such as race, gender, and socioeconomic status.

Verb Tense

The literature assumes that past tense narrative is the norm for truthful accounts of past events (Dulaney, 1982; Sapir, 1987; Rudacille, 1994). However, as Porter and Yuille (1996) demonstrate, it is deviations from the past tense that correlate with deception. Indeed, changes in tense are often more indicative of deception than the overall choice of tense. The most often cited example of tense change in a criminal statement is that of Susan Smith, who released the brake on her car letting her two small children inside plunge to their deaths. "I just feel hopeless," she said. "I can't do enough. My children wanted me. They needed me. And now I can't help them. I just feel like such a failure." While her statements about herself were couched in the present tense, those about her children were already in the past.

Hedges

The terms 'hedge' and 'hedging' were introduced by Lakoff (1972) to describe words "whose meaning implicitly involves fuzziness", e.g., *maybe, I guess,* and *sort of.* The use of hedges has been widely studied in logic and pragmatics, and for practical applications like translation and language teaching (for a review, see Schröder & Zimmer, 1997). In the forensic psychology literature, it has been correlated with deception (Knapp et al., 1974; Porter & Yuille, 1996; Vrij & Heaven, 1999).

Hedge types in our data include non-factive verbs like *think* and *believe*, non-factive NPs like *my understanding* and *my recollection*, epistemic adjectives and adverbs like *possible* and *approximately*, indefinite NPs like *something* and *stuff*, and miscellaneous phrases like *a glimpse* and *between 9 and 9:30.*

The particular types of hedging that appear in our data depend heavily on the socioeconomic status of the speaker and the type of crime. The 285 hedges in Jeffrey Skilling's 7562 word Enron testimony include 21 cases of *my recollection*, 9 of *my understanding*, and 7 of *to my knowledge* while the 42 hedges in the car thief's 2282 word testimony include 6 cases of *shit* (*doing a little painting, and roofing, and shit*), 6 of *just* and 4 of *probably*. Despite the differences in style, however, the deceptive behavior in both cases is similar.

Changes in Referential Expressions

Laboratory studies of deception have found that deceivers tend to use fewer self-referencing expressions (*I, my, mine*) than truth-tellers and fewer references to others (Knapp et al., 1974; Dulaney, 1982; Newman et al., 2003). In examining a specific real world narrative, however, it is impossible to tell what a narrator's truthful baseline use of referential expressions is, so the laboratory findings are hard to carry over to actual criminal narratives.

On the other hand, changes in the use of referential expressions, like changes in verb tense, have also been cited as indicative of deception (Sapir, 1987; Adams, 1996), and these changes can be captured formally. Such changes in reference often involve the distancing of an item; for example, in the narrative of Captain McDonald, he describes 'my wife' and 'my daughter' sleeping, but he reports the crime to an emergency number as follows, with his wife and daughter referred to as *some people*:

> So I told him that I needed a doctor and an ambulance and that *some people* had been stabbed.

Deceptive statements may also omit references entirely. Scott Peterson's initial police interview is characterized by a high number of omitted first person references:

> BROCCHINI: You drive straight home?
> PETERSON: To the warehouse, dropped off the boat.

4.2 Identifying a Text Passage as Deceptive or Non-deceptive

The presence or absence of a cue is not in itself sufficient to determine whether the language is deceptive or truthful. Linguistic hedges and other deception indicators often occur in normal language use. We hypothesize, however, that the distribution and density of the indicators would correlate with deceptive behavior.[2] Areas of a narrative that contain a clustering of deceptive material may consist of outright lies or they may be evasive or misleading, while areas lacking in indicator clusters are likely to be truthful.

We use a moving average (MA) program to find clusters of indicators in a text. Initially, the MA assigns each word in the text a proximity score based on its distance, measured in word count, to the nearest deception indicator. Each score is then recalculated by applying a MA window of N words. The MA sums the scores for N/2 words to the left and right of the current word and divides the result by N to obtain the revised score. Clusters of low word scores indicate deceptive areas of the text, high scoring clusters indicate truthful areas. Hence, when applied to a text, the MA allows us to segment an entire text automatically into non-overlapping regions that are identified as likely true, likely deceptive or somewhere in between.

Our approach assumes that the input text will contain sufficient language to display scoring patterns. This rules out, for example, polygraph tests where answers are confined to Yes or No as

[2] Currently the density algorithm does not take into account the possibility that some indicators may be more important than others. We plan to use the results of this initial test to determine the relative contribution of each tag type to the accuracy of the identification of deception.

well as short answer interviews that focus on simple factual statements such as names and addresses. Based on the data examined so far, we estimate the analysis requires a minimum 100 words to produce useful results.

5. Corpora and Annotation

The corpus used for developing our approach to deception detection was assembled from criminal statements, police interrogations, depositions and legal testimony; the texts describe a mix of violent and property crimes, white collar crime and civil litigation. Because of the difficulty in obtaining corpora and ground truth information, the total corpus size is small--slightly over 30,000 words.

For this experiment, we selected a corpus subset of 25,687 words. Table 1 summarizes the corpus subset:

Source	**Word Count**
Criminal statements (3)	1,527
Police interrogations (2)	3,922
Tobacco lawsuit deposition	12,762
Enron congress. testimony	7,476
Total	25,687

Table 1. Corpora Used in the Experiment

Each document in the experimental corpus was tagged for two factors: (1) linguistic deception indicators marked words and phrases associated with deception, and (2) True/False tags marked propositions that were externally verified.

5.1. Linguistic Annotation (Tagging)

A team of linguists tagged the corpus for the twelve linguistic indicators of deception described above. For each document in the corpus, two people assigned the deception tags independently. Differences in tagging were then adjudicated by the two taggers and a third linguist. Because the original tagging work was focused on research and discovery, inter-rater reliability statistics are not very revealing. However, current work on new corpora more closely resembles other tagging tasks. In this case we have found inter-rater reliability at 96%.

Tagging decisions were guided by a tagging manual that we developed. The manual provides extensive descriptions and examples of each tag type. Taggers did not have access to ground truth facts that could have influenced their tag assignments.

5.2. True/False Annotation

We then examined separate copies of each narrative for propositions that could be externally verified. The following is a single proposition that asserts, despite its length, one verifiable claim—the birthrate went down:

> *The number of births peaked in about 1955 and from there on each year there were fewer births. As a result of that each year after 1973 fewer people turned 18 so the company could no longer rely on this tremendous number of baby boomers reaching smoking age.*

Only propositions that could be verified were used. Verification came from supporting material such as police reports and court documents and from statements internal to the narrative, e.g. a confession at the end of an interview could be used to support or refute specific claims within the interview. The initial verification tagging was done by technical and legal researchers on the project. The T/F tags were later reviewed by at least one other technical researcher.

The experimental corpus contains 275 verifiable propositions. Table 2 gives examples of verified propositions in the corpus.

Example	**True**	**False**
I didn't do work specifically on teenage smoking		√
All right, man, I did it, the damage	√	
Black male wearing a coat.		√

Table 2. Examples of Verified Propositions

6. Results

The dataset contained 275 propositions, of which 164, or 59.6%, were externally verified as False and the remainder verified as True. We tested the ability of the model to predict T/F using Classification and Regression Tree (CART) analysis (Breiman, et al. 1984)[3] with 25-fold cross-validation and a misclassification cost that penalizes True misclassified as False. Table 3 shows the results of the CART analysis:

[3] We used the QUEST program described in Loh and Shih (1997) for the modeling. QUEST is available at http://www.stat.wisc.edu/~loh/quest.html.

Actual Class	Predicted Class			
		False	True	% Correct
	False	124	40	75.6
	True	29	82	73.8

Table 3. T/F Classification Based on Cue Density

We can conclude that the model identifies deceptive language at a rate significantly better than chance. Moreover, by tuning the scores to favor high recall for false propositions, it becomes possible to adapt the model to applications where low precision on true propositions is not a drawback, e.g. pre-trial interviews where investigators are looking for leads. The results in Table 4 show how we might gear the analysis to this class of applications.

Actual Class	Predicted Class			
		False	True	% Correct
	False	151	13	92.6
	True	66	45	40.5

Table 4. Penalizing F Misclassified as T

Finally, it should be noted that input to the analysis consisted of individual files with some files marked for topic changes. In preparing the data for this test, we found that, in many cases, the moving average allowed the low scores assigned to deceptive language to influence the scores of nearby truthful language. This typically occurs when the narrative contains a change in topic. For example, in the deposition excerpt below, there is a topic change from teenage smokers to the definition of psychographic studies. The hedge *so far as I know* belongs with the first topic but not the second. However, the moving average allows the low scores triggered by the hedge to improperly affect scores in the new topic:

> Q: *Do you know anybody who did have data that would allow a market penetration study of the type I've asked about to be performed.*
> A: *{So far as I know%HEDGE} only the federal government.*
> Q: *Are you familiar with the phrase psychographic study from your work at Philip Morris?*
> A: *Yes.*
> Q: *What is a psychographic study?*

To mitigate the effect of topic change, we inserted eleven topic change boundaries. The results suggest that language is "reset" when a new topic is introduced by the interviewer or interviewee.

7. A Deception Indicator Tagger

The results described in the previous section provide support for the deception indicator (DI) approach we have developed. For the implementation, we selected a subset of tags whose contextual conditions were well established by the literature and our own investigation. In these cases we were able to formalize the rules for automatic assignment of the tags. We excluded tags whose contextual conditions are still being researched, i.e., tag assignments that require human judgment.

The tagger was constructed as a rule-based system that uses a combination of context-free and context sensitive substitutions. An example of a context free substitution is “Mark all occurrences of *Oh, God* as an overzealous statement”. A context sensitive substitution is the rule that interprets *something* as a hedge if it is not modified, i.e., followed by a relative clause or prepositional phrase.

In some cases the tagger refers to structure and part of speech. For example, *may* as a modal verb (*may_MD)* is a hedge. Certain verb+ infinitive complement constructions, e.g. *I attempted to open the door,* make up a qualified assertion. Syntactic structure is assigned by the CASS chunk parser (Abney, 1990). Part of speech tags are assigned by Brill’s tagger (Brill, 1992). The DI tag rules apply to the output of the parser and POS tagger.

The subset of tags implemented in the tagger comprises 86% of all tags that occur in the training corpus. To see how well the DI tagger covered the subset, we first ran the tagger on the training corpus. 70% of the subset tags were correctly identified in that corpus, with 76% precision. We then tested the tagger on a test corpus of three files. Each file was also handtagged by linguistic researchers on this project. The results of the test are given in Table 5. Tag amounts refer to the number of tags belonging to the subset that was implemented.

File name	Handtags	Autotags	Correct Tags
confession	31	20	19
peterson	186	160	108
deposition	720	665	625
Total	937	845	752

Table 5. DI Tagger Results on Three Test Files

Table 6 provides a summary of the tagger's performance.

File name	Recall	Precision
confession	.61	.95
peterson	.58	.675
deposition	.868	.939
Average	.686	.853

Table 6. Summary of DI Tagger Results

These results may reflect a bias in our training data towards legal testimony—depositions are strongly represented in the corpus, police and criminal data less so. Our test corpus consists of a police interview ('peterson'), a criminal statement ('confession') and a deposition ('deposition'). The tagger's best performance is associated with the deposition.

8. Conclusion

This paper has presented new results in the study of language-based cues to deception and truthfulness; these results come entirely from "real world" sources—criminal narratives, interrogations, and legal testimony. Our goal is to provide a method of evaluating declarations within a single narrative or document rather than deeming an entire narrative (or narrator) as truthful or deceptive.

We first compared the predictions of linguistic cues that we adapted from the literature on deception against actual True/False values that were manually determined for 275 propositions in our corpus. Predictions from the linguistic indicators were determined by scoring the density of indicators in text areas that contain the propositions and using classification and regression to determine cut-off values for truth probabilities.

We then evaluated the performance of an automated tagger that implements a large subset of the linguistic indicators verified in our first experiment. The automated tagger performed well on test data, averaging 80.2% correct when compared with human performance on the same data.

The results strongly suggest that linguistic cues provide a guide to deceptive areas of a text. The predictions based on linguistic cues were correct in distinguishing False propositions over 75% of the time, and over 90% for applications where recall of False, but not True, is required. Results of the automatic tagger's performance suggest that we will eventually achieve a fully automated system for processing depositions and other documents in which veracity is an important issue.

References

Abney, S. 1990. Rapid incremental parsing with repair. In *Proceedings of the 6th New OED Conference: Electronic Text Research*, pp. 1-9. University of Waterloo, Waterloo, Ontario.

Adams, S. 1996. Statement analysis: What do suspects words really reveal? *The FBI Law Enforcement Bulletin*. 65(10). www.fbi.gov/publications/leb/1996/oct964.txt

Adams, S. 2002. *Communication under stress: indicators of veracity and deception in written narratives.* Ph.D. dissertation, Virginia Polytechnic Institute and State University

Brill, E. 1992. A simple rule-based part-of-speech tagger. In *Proceedings of the Third Conference on Applied Natural Language Processing,* pp. 152-155. Trento, Italy.

DePaulo, B. M., J.J. Lindsay, B.E. Malone, L. Muhlenbruck, K. Charlton, and H. Cooper. 2003. Cues to deception. *Psychological Bulletin*, 129(1), 74-118.

Dulaney, E.F. Jr. 1982. Changes in language behavior as a function of veracity. *Human Communication Research* 9, 75-82.

Hirschberg, J., S. Benus, J. Brenier, F. Enos, S. Friedman, S. Gilman, C. Girand, M. Graciarena, A. Kathol, L. Michaelis, B. Pellom, E. Shriberg and A. Stolcke. 2005. *INTERSPEECH 2005.* Sept. 408, Lisbon, Portugal.

Knapp, M.L., Hart, R.P., and Dennis, H.S. 1974. An exploration of deception as a communication construct. *Human Communication Research, 1,* 15-29.

Lakoff, G. 1972. Hedges: A study in meaning criteria and the logic of fuzzy concepts. In

Papers from the 8th Regional Meeting, Chicago Linguistic Society.

Loh, W.-Y. and Shih, Y.-S. 1997. Split selection methods for classification trees. *Statistica Sinica* 7:815-840.

Miller, G. R. and J. B. Stiff. 1993. *Deceptive Communication.* Sage Publications. Thousand Oaks, CA.

Newman, M. L., Pennebaker, J. W., Berry, D. S. and J. M. Richards. 2003. Lying words: predicting deception from linguistic styles. *Personality and Social Psychology Bulletin.* **29**, 665-675.

Porter, S. & Yuille, J. (1996). The language of deceit: An investigation of the verbal clues in the interrogation context. *Law & Human Behavior,* 20(4) 443-458.

Rudacille, W.C. 1994. *Identifying Lies in Disguise*. Kendall Hunt. Dubuque, IO.

Sapir, A. 1987. *Scientific Content Analysis (SCAN).* Laboratory of Scientific Interrogation. Phoenix, AZ.

Sapir, A. 1995. *The View Guidebook: Verbal Inquiry – the Effective Witness*. Laboratory of Scientific Interrogation. Phoenix, AZ.

Schröder, H. and D. Zimmer. 1997. Hedging research in pragmatics: A bibliographical research guide to hedging. In R. Markkanen and H. Schroder (eds.) *Hedging and Discourse: Approaches to the Analysis of a Pragmatic Phenomenon in Academic Text.* Walter de Gruyter, Berlin.

Shuy, R. 1998. *The Language of Confession, Interrogation and Deception*. Sage Publications, Thousand Oaks, CA.

Smith, N. 2001. Reading between the lines: An evaluation of the scientific content analysis technique (SCAN). Police Research Series. London,UK.
www.homeoffice.gov.uk/rds/prgpdfs/prs135.pdf

Steller, M. and G. Kohnken. 1989. Criteria-Based Content Analysis. In D.C. Raskin (ed.) *Psychological Methods in Criminal Investigation and Evidence*. Springer-Verlag, New York, 217-245.

Undeutsch, U. 1989. The development of statement reality analysis. In J.C. Yuille (ed.) *Credibility Assessment.* Dordrecht: Kluwer, 101-121.

Undeutsch, U. (1954). Die Entwicklung der gerichtspsychologischen Gutachtertatigkeit. In A. Wellek (Ed.), Bericht uber den 19, Kongress der Deutschen Gesellschaft fur Psychologie (pp. 1132-154). Gottingen: Verlag fur Psychologie.

Vrij, A. 2000. *Detecting Lies and Deceit*. John Wiley & Sons, Chichester, UK.

Vrij, A. and Heaven, S. 1999. Vocal and verbal indicators of deception as a function of lie complexity. *Psychology, Crime, and Law 5*, 203-215.