COMPUTERIZED LINGUISTICS AND PSYCHOLOGY - OVERCOMING THE POLYGRAPH'S DRAWBACKS

Dan OPHIR, Ph.D.
Afeka Academic College of Engineering, Israel
218 Bney Efraim Str. Tel Aviv, 69107 Israel
dano@afeka.ac.il

Abstract: We developed an auxiliary tool, named "Software Human Reliability Estimator" (SHRE), which in certain cases can replace the polygraph. The polygraph is not always effective in measuring the reliability of a witness. For instance, the polygraph is ineffective when the witness believes that the testimony is the truth even when in reality it is not. In such cases, an alternative objective test is required. Another disadvantage of the polygraph test lies in the lack of discreetness owing to the requirement that the witness must agree to undergo a polygraph test. In addition, the polygraph test cannot be performed in real time because of its cumbersome and bulky nature.

These drawbacks have motivated the search for alternatives to the polygraph. Herein, we suggest a methodology accompanied by a corresponding software package that overcomes the mentioned shortcomings of the polygraph.

The methodology is based on a computer-assisted cognitive behavioral therapy methodology (CBT) (Burns 1999). CBT was originally developed for psychological treatment and can be used to characterize personalities. This methodology can also be used to find the individual's personality disturbances and to evaluate the reliability of a witness. The CBT methodology assumes that the cognitive thoughts of a human are expressed in his language. In the literature, about ten categories of thoughts are determined, and so called ***distorted thoughts***, indicate a behavioral deviation. Based on the above assumption, it is possible to map thoughts, including ***distorted thoughts*** and analyze them methodically with the help of linguistic tools. These tools should be able to scan the mapping and discover ***distorted thoughts*** as classified by the CBT method.

We will use extreme situations as examples to illustrate ***distorted thoughts***. The mentioned situations will refer to time description (always, never), location (everywhere, nowhere), quantity (everything, nothing, nobody), possibility (must, forced, incapable) etc. These types of expressions leave no doubt as to their meanings.

The linguistic analysis is performed at two levels: semantic and syntactic. The first stage is the semantic analysis. Here, the vocabulary of the sentence is analyzed.

The known linguistic term, *quantitative-semantics*, is given a special significance since it enables a pre-ranking of the nouns, adjectives, adverbs beyond their regular usage. *Quantitative semantics* analysis searches especially for superlatives such as "*never*", which indicate an extreme case. This analysis is supported in the first stage by using an expression named "*distinguished*".

In order to find *distinguished* expressions it is recommended to use in the second stage of the analysis a methodology borrowed from *formal-languages*, a field in computer sciences. This analysis is supposed to strengthen or eliminate the indications found in the first analysis stage, the semantic analysis.

SKOMPUTERYZOWANA LINGWISTYKA I PSYCHOLOGIA POKONUJĄC OGRANICZENIA POLIGRAFU

Abstrakt: Artykuł opisuje urządzenie, które w niektórych przypadkach zastępuje poligraf. Poligraf niekiedy jest mało efektywny w mierzeniu wiarygodności świadka, np. poligraf nie jest skuteczny, gdy świadek wierzy w swoje zeznanie, nawet jeśli w rzeczywistości nie jest ono prawdziwe. W takich przypadkach pomocny jest alternatywny test wiarygodności.

Inną słabą stroną poligrafu jest to, że nie można wykonać testu bez zgody świadka. Poza tym z powodu jego uciążliwości, test na poligrafie nie może być wykonany *in vivo*. Te wady inspirowały do poszukiwania alternatyw dla poligrafu. W niniejszym artykule sugerujemy metodologię i towarzyszące jej oprogramowanie, które przezwyciężają wymienione niedogodności poligrafu.

Metodologia jest oparta na kognitywnej metodzie terapii behawioralnej dla komputera (CBT) (Burns, 1999). CBT była przede wszystkim przeznaczona do terapii psychologicznej, ale może też być używana do charakteryzowania osobowości. Ta metodologia może być również stosowana do określania zaburzeń osobowości świadka i oszacowania stopnia jego wiarygodności.

Lingwistyczna analiza jest wykonywana na dwóch poziomach: semantycznym i syntaktycznym. Pierwszy etap stanowi analiza lingwistyczna, która bada słownictwo.

Nowe podejście lingwistyczne polega na analizie *semantyki ilościowej*, która określa pewien typ skalowania, przede wszystkim przymiotników i przysłówków w potocznym użyciu.

Analiza *semantyki ilościowej* poszukuje głównie superlatywów (np. 'nigdy'), które wskazują skrajne przypadki. Pierwszy etap tej analizy polega na określeniu odpowiedniego użycia tzw. „słów wyróżnionych" (przedstawionych na ilustracji 1). Aby znaleźć te „wyróżnione wyrażenia", na drugim etapie badania należy dokonać analizy syntaktycznej (ilustracje 5 i 6). Ten proces powinien potwierdzić wyniki otrzymane na pierwszym etapie badań.

Komputeryzując analizę składni zdania, możemy szybko zakwalifikować tekst, używając metod *semantyki ilościowej*, pokazanej poniżej. Skomputeryzowana semantyka ilościowa może być stosowana jako narzędzie pomocnicze w psychologii i autopsychoterapii.

Kognitywna terapia behawioralna(CBT) jest obecnie bardzo popularną metodą wśród psychoterapeutów i stała się bodźcem dla rozwoju *semantyki ilościowej*. Jej główna zaleta polega na prostocie i schematycznej metodologii. Te cechy ułatwiły stworzenie i wdrożenie skomputeryzowanego modelu kognitywnego terapii(ICBT). ICBT jest skomputeryzowanym CBT i zajmuje się obróbką informacji. Istotną cechą tego modelu jest znajdowanie tzw. zniekształconych myśli. Myśl zniekształcona jest myślą, która próbuje przedstawić rzeczywistość, ale w zdeformowanej formie (Burns, David D., 1999).

Na przykład następujące przypadki przedstawiają myśl zniekształconą: Student, który otrzymał notę dostateczną, konkluduje: *Jestem kompletnym durniem.* Chłopak po kłótni ze swoją dziewczyną wyciąga wniosek: *Dziewczyny wszystko psują. Ja nigdy nie będę w stanie dojść do ładu z dziewczynami.*

Myśli kognitywne są sformalizowane przez ludzki mózg w języku naturalnym w formie mówionej lub pisanej. Z tego powodu analiza każdej myśli może być wykonana na zdaniu – lingwistycznym odpowiedniku myśli. Tekst zawierający zniekształcone myśli charakteryzuje się występowaniem superlatywów.

בלשנות ממוחשבת ופסיכולוגיה – התגברות על חסרונות הפוליגרף

Hebrew Abstract תקציר: הפוליגרף יעיל **במידה ידועה במדידת** אמינותו של עד, אבל הוא אינו יעיל כאשר העד מתנהג כאילו העדות היא האמת, **גם כשאיננה כך**. חסרונות **נוספים** של **בדיקת** פוליגרף נעוצים בהעדר דיסקרטיות הנובעת מהצורך לקבל את הסכמתו של העד לביצוע בדיקת הפוליגרף. **כמו כן**, בדיקת פוליגרף אינה יכולה להתבצע בזמן אמת, בגלל הסרבול והעומס **הכרוכים בכך**. חסרונות אלו של הפוליגרף **עוררו את השאיפה למצוא תחליף** לפוליגרף. המאמר הזה **מציג** כלי עזר שעשוי להחליף את הפוליגרף במקרים מסוימים. כלי זה מציע מתודולוגיה המלווה בחבילת תוכנה המתגברת על החסרונות הנ"ל של הפוליגרף. הכלי המוצע **מכונה "*תוכנת מעריך אמינות אנושית*":**

Software for Human Reliability Estimator – SHRE). המתודולוגיה מבוססת על שיטת טיפול קוגניטיבי-התנהגותי הנעזרת במחשב **(computer-assisted cognitive behavioral therapy methodology**, או **CBT**, ברנס 1999). שיטת ה-**CBT**, שפותחה במקור לטיפול פסיכולוגי, יכולה לשמש גם לאפיון אישיות ולמציאת הפרעות אישיות של העד ולהערכת אמינותו**.** שיטת ה-**CBT** יוצאת מתוך ההנחה כי מחשבותיו (הקוגניטיביות) של האדם באות לידי ביטוי בהתבטאותו המילולית (או במבעיו הלשוניים). בספרות אובחנו כעשרה סוגי מחשבות, המכונים *עיוותי מחשבות*, שמצביעים על או **משקפים** סטייה התנהגותית. על סמך הנחת היסוד שצוינה, ניתן למפות את מחשבות האדם, כולל *עיוותי מחשבה*, בכתב או **באמצעי אחר**, ולנתח אותן באופן שיטתי בעזרת כלים בלשניים. הכלים הבלשניים אמורים לסרוק את מיפוי המחשבות ולגלות את *עיוותי המחשבה* שמוינו על ידי חלוצי שיטת ה-**CBT** (במידה שהם קיימים). להמחשת *עיוותי מחשבה* מוצגות **במאמר זה** דוגמאות למחשבות המשקפות מצבים קיצוניים כגון תיאורי זמן (תמיד, אף פעם), מקום (בכל מקום, בשום מקום), כמות (כל, כולם, אף אחד), יכולת (מוכרח, לא מסוגל) ועוד, כלומר, ביטויים שאינם משאירים מקום לספק **בדבר קיצוניותם**.

הניתוח הבלשני הזה מבוצע בשתי רמות: סמנטית וסינטקטית. השלב הראשון הינו ניתוח סמנטי הבודק את אוצר המילים (של המשפט). המונח הבלשני "*סמנטיקה-כמותית*" משמש כאן במשמעות מיוחדת**:** *סמנטיקה-כמותית* מאפשרת דירוג ראשוני של כל שמות **ה**עצם, שמות **ה**תואר, ותארי **ה**פועל מעבר לשימושם הרגיל. ניתוח של *סמנטיקה-כמותית* מחפש במיוחד סופרלטיבים כגון "*אף פעם*", אשר מצביעים על מקרים קיצוניים. ניתוח זה מבוסס בשלב ראשון על השימוש בביטויים הקרויים "*מיוחדים*, **יוצאי דופן**". כדי למצוא את הביטויים *המיוחדים* בשלב השני של הניתוח, אנו משתמשים בשיטת ניתוח תחבירי השאולה מתחום השפות הפורמאליות **ב**מדעי המחשב. ניתוח זה אמור לאשש או לשלול את הממצאים שזוהו בצעד הראשון של הניתוח, כלומר בשלב הניתוח הסמנטי.

Introduction

The present paper is interdisciplinary in its functions and applications. The computerized linguistic analysis presented here can be used, for example, in the disciplines of law and psychology.

Law, linguistics, and their connection to computers have been previously studied (Cotterill 1968; Gibbons 2003; Shuy 1966; Olson 2004). The present article focuses on linguistic analysis of cross-examination texts (Colma 1970, Salhany 2006, Glisan 1991). Part of this paper is about the use of semantic analysis in psychology in that the vocabulary of thoughts is checked. ***Distorted thought*** text is characterized by the use of superlatives such as "***never***". The definitions of ***distorted thought*** has been defined and categorized by the developers of CBT (Burns 1999 and Greenberger and Padesky 1995). These categorizations can be used to automatically recognize and classify written statements by a computerized analysis (Kearns 2000; Knuth 1964). This analysis is based, in the first iteration, on the corresponding use of expressions called '***distinguished***' words. To find these *distinguished* terms, which possibly indicate cognitive distortions, quantitative semantics is introduced.

The present article's analysis is based on the title's two main components, namely linguistics and psychology. Linguistics' two main branches, semantics and syntax, were used in the development of a software tool called Software Human Reliability Estimator (SHRE). SHRE can be used as an alternative to a polygraph. The syntax's extended treatment is represented by the following elements: BNF definitions (a computer sciences method using formal languages for defining a computer languages' syntax), speech parts decomposition and parsing tree construction.

Both of the linguistics parts complement each other and they form a validation of the results. Further analysis is done using the program's psychological aspects which estimates the reliability of the individual and the results obtained.

The psychological section, called ***'Evaluation',*** recognizes a cognitive distortion and if required, replaces it by a proposed correction. The correction of the ***distorted***

thought isn't generally the purpose of the SHRE, but some of the software's applications may be in guiding the user towards self-improvement.

Semantics: Quantitative semantics preface

The linguistic term ***quantitative-semantics*** is widely used in the context of the introduced application***. Quantitative-semantics*** allows for a type of scaling of primarily adjectives and adverbs beyond their common usage.

Quantitative-semantics analysis looks in particular for superlatives such as '***never'***, which hints at an extreme case. The analysis is based, in the first iteration, on the corresponding use of expressions termed '***distinguished'*** words, representing distortion thoughts. To find these ***distinguished*** terms, one must analyze a syntactic sentence. This should reconfirm the indications iterated in the first stage.

Using computerized sentence analysis, we can quickly classify texts using the semantic-quantification methods shown below. Computerized quantitative-semantics can be used as an auxiliary tool for psychology and for self-psychotherapy.

Cognitive behavioral therapy (CBT) is currently a very popular method among psychotherapists and was the impetus for developing ***quantitative semantics.*** Its main advantage lies in its simplicity and in its schematic methodology. These characteristics facilitate the creation of a computer-implemented cognitive therapy model, iCBT ("i" stands for information). iCBT is a computerized CBT that deals with an auxiliary computerized ***information*** processing,

The crux of this model lies in finding the person's so-called '***distorted thought'***. A ***distorted thought*** is a thought that tries to represent reality, but instead gives a distorted or unrealistic result (Burns 1999). For example, the following sentence represents a ***distorted thought***: A student who has received a grade C concludes "***I am a complete moron***…". Another example deals with someone who, after quarreling with his girlfriend, concludes: "***Girls always spoil the relationship. I will never be able to hold on to a girlfriend***…"

Cognitive thoughts may be formulated by the human brain into a natural language, namely, into meaningful spoken or written sentences. Therefore, the analysis of any thought is actually performed on the sentence, the thought's linguistic counterpart. ***Distorted thought*** text is characterized by the use of superlatives. In order to help to find them in an analyzed text, the following additional general terms are introduced for better understanding the further formulations.

a. *Human Factor* – the aspects that deal with behavioral sciences, namely, psychology and more specifically, CBT – cognitive behavioral therapy.
b. *Languages* – a natural language is the interface between subconscious thoughts and conscious speech.
c. *Measurement* – the measurements are performed to evaluate the tested text. The ***distorted thought*** is then transformed into its normative counterpart.

Semantics: Quantitative semantics – detailed description

Semantics (Kearns 2000;Ferdinand and Harri 1986) is a linguistic area of study that tries to parse the significance of a sentence and its parts. This is one of the required fields in cognitive thought analysis. In order to better serve the needs of iCBT, we will propose some ranking of the vocabulary.

This ranking is termed hereafter as ***Quantitative-Semantics***. It is defined as follows: The parts of speech, for example, adjectives, adverbs, and nouns are organized into "family-groups" containing sorted members in the family group. Each group treats some property represented by an abstract noun (or term) such as ***speed***, ***hunger, and feeling (hot, cold, etc.).***

The sorting is done according to the intensity of the meaning of the word in the family-group, starting from a lower intensity, proceeding through moderate words and then to the higher ones. The members of the family-group appear with their attached intensity value.

For example, the sequence of the following words represents the idea: ***{{cold, -2}, {hypothermal,-1}, {lukewarm, 0, {tepid, 0}, {warm,1},{ hot,2}}.***

This ranking will first be presented in BNF notation (Figure 1) with the auxiliary notation (Figure 2), and then its usage will be analyzed.

This is only a partial list of a much longer one that is being created to indicate contrast.

Syntax: Defining Bacchus normal form or Bacchus-Naur form (BNF) notation

The BNF method for describing the linguistic characteristics of various ***distorted thoughts*** is widely used in defining the syntax of programming languages (Knuth 1964). This technique will be used in the context of iCBT. Understanding this technique is essential for further understanding this article.

Figure 1 (a): Defining a Grammar, which can generate the extreme terms, so called "*distinguished*" words indicating ***distorted-thoughts:*** main set.

1. *<determining-term>::=<extreme-term> | <moderate-term>*
2. *<extreme-term> ::= <minimal-term> | <maximal-term>*
3. *<minimal-term>::= <minimal-timing-term> | <minimal-location-term> | minimal-personal-term> | <minimal-still-term>*
4. *<maximal-term> ::= <maximal-timing-term> | <maximal-location-term> | maximal-personal-term> | <maximal-still-term>*
5. *<moderate-term>::= <moderate-timing-term> | <moderate-location-term> | moderate-personal-term> | <moderate-still-term>*
6. *<emotional-term>::= < negative-emotional-term > | < positive-emotional-term>*
7. *< negative-emotional-term >::= sadness | unhappiness |despondency | depressing | anxiety | restlessness | unease | dissatisfaction | discontent*
8. *< positive-emotional-term>::= happiness | calmness | satisfaction | contentment*

9. *<minimal-timing-term>::= never | not at all | not at any time | not ever | not in any way*
10. *<maximal-timing-term>::= always | constantly | without stopping | anytime*
11. *<moderate-timing=term>::= sometimes | once in awhile | occasionally | often | seldom | frequently | in many cases*

12. *<minimal-location-term>::= nowhere | not anywhere | in no place*
13. *<maximal-location-term>::= everywhere | in every place | in every location*
14. *<moderate-location-term>::= here and there | somewhere | someplace | some location | any place*

(a)

Figure 1 (b): (continuation)

1. *<minimal-manner-term> ::= wrong | behave unjustly toward | injure | harm | violate | malign | discredit*
2. *<maximal-manner-term>::= well | excellently | in a good manner | appropriately |properly | significantly | in good spirit | completely |totally*
3. *<positive-relation>::= love | like | polite |respect | honor |dignity | face |glory | homage| honorableness | kudos | regard | reputation | comity | success*
4. *<negative-relation>::= hate | cruel | scorn | contempt | derision |disregard | disparagement| flippancy | levity | failure*
5. *<moderate-relation>::= acquaintance | friend*
6. *<sharp-conscientiousness>::= should | must | have | mandatory | obligatory*
7. *<moderate-conscientiousness>::= may | might | can | could | maybe | possible | probable*
8. *<negative-label-affront>::= stupid | fool | pig | monkey | donkey*
9. *<positive-label-affront>::= wise | smart | talented | lion | cat*

(b)

Figure 2: Definitions of auxiliary terms

<connector>::= <cause-connector> | <pointing-connector>
<cause-connector> ::= therefore | because | and so | hence | namely | i.e.| however
<pointing-connector>::= that | which | who
<importance-term>::= <neglecting-term> | <emphasizing-term>
<less-important-term>::= by coincidence | it's nothing | good luck | only
<more-important-term>::= intentional | knowingness | willfulness

The BNF methodology is based on using symbols (Figure 3) for its notation (Chomsky 1957).

The ***sharp-brackets*** contain the terms to be defined and the terms that already have been defined. The set of two colons and the equal operator define the assignment operator in *BNF* notation. **"|"** denotes the alternative operator, as shown in the example

(Figure 4). The ***rectangle-brackets*** denote an option. The ***tilde*** operator denotes a negation or complementation. The ***space*** denotes the concatenation operator and the regular parentheses control the precedence of the operators as they do in algebra (Figure 4). The three dots denote a repetition.

Figure 3: BNF's conventional symbols

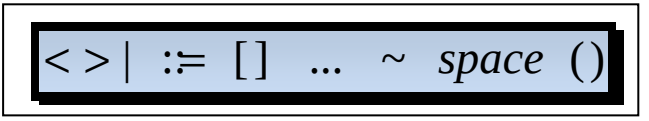
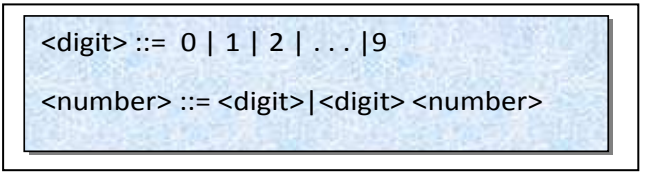


Figure 4: An example using the BNF symbols to define the term *number*.



Syntax: Analysis

(i) Syntactic parts

For the computer to accurately analyze a sentence, the sentence must be decomposed into its syntactic parts. Assuming that the sentence's words are found in a dictionary along with their corresponding part of speech, e.g., an adjective, noun, verb, and adverb, it is possible to classify the sentence's words into their syntactic role within the sentence. Knowing the syntactic function of the sentence's words such as the subject, object, and predicate will help analyze the semantics or the meaning of the sentence. This meaning enables us to automatically recognize ***distorted thoughts***.

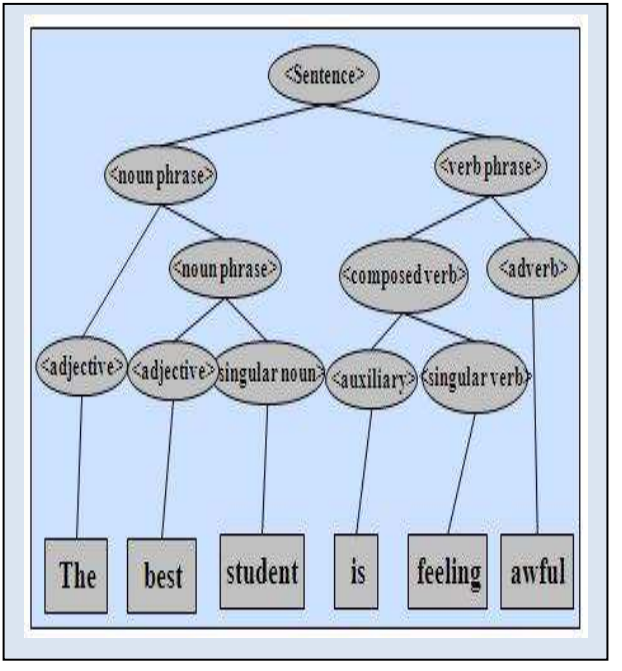
(ii) Examples

A simple sample sentence "***The best student is feeling awful"*** will be analyzed. Initially, the BNF corresponding rules are applied **(**Figure 5**)** and the corresponding derivation tree (Figure 6) is obtained using the ***syntactic-structure*** method. The usefulness of the BNF notation and of decomposing the syntactic structure will be further discussed.

Figure 5: Derivation – BNF rules of a sentence: ***"The best student is feeling awful."***

```
<sentence> ::= <noun phrase> <verb phrase>
<noun phrase>   ::= <adjective> <noun-phrase>|
          <adjective><singular noun>
<verb phrase> ::= <verb> <adverb>
<verb>  ::= <singular verb>| <composed verb>
<composed verb> ::= <auxiliary> <verb>
<auxiliary> ::=  is | are | have
<adjective> ::=  the | best
<noun> ::=  student
<singular verb> ::= feeling
<adverb> ::=  awful
```

Figure 6: The derivation/parsing tree of the sentence, ***"The best student is feeling awful",*** uses the rules from Figure 5.



A more extensive example is given by a sample of a cross-examination transcript (Figure 7), (Salhany R. 2006, 86-87). The background story of the interrogation in Figures 7(a-b) is as follows:

Alfred Rouse was prosecuted for the murder of an unknown man. His counsel, Donald Fennimore, led him through the various lies he made and asked him to explain each of the lies. The counsel's obvious purpose was to lessen the impact of any cross-examination as to why he had lied. This is similar to defense counsels leading defendants through their criminal records to dampen their effect on the jury before the prosecutor has a chance to raise any part of the record against the defendant.

Norman Burkett, who prosecuted Rouse, decided directly to raise the issue of those lies in his first questions.

Even with the advance preparation of the defendant by the defense counsel, the counsel cannot foresee the prosecutor's questions and therefore the defendant is forced to improvise answers and then falls in the trap laid out by the prosecutor who uses sophisticated questions. The answers of the defendant will be analyzed by the methodology introduced next.

The shown cross-examination (Figure 7) illustrates the witness's ***distorted-thought*** through his use of expressions such as ***always***, ***never, which*** are ***extreme-time*** expressions, indicating a "***minimization and magnification***" distortion type (Figure 1). The presented technique is even more effective in analyzing the ***character-evidence,*** which is composed of longer texts with fewer interruptions.

Figure 7(a): A transcript of a cross-examination, in which our analysis may improve witness evaluation (Glisan, James Lindsay. 1991 p. 86-87).

Q. Rouse, when you told my learned friend that the lies you had told in Wales were unfortunate, what did you mean?

A. Well, I think that it is always best — I have always been noted for telling the truth for the whole of my life; I am not used to telling lies. At the time I thought it was the best thing to do.

Q. Why? Why was lying better than telling the truth?

A. Because there are many members of the family, for one thing, and I should have to tell the story over and over again, and I did not like to tell it with ladies present.

Q. Why tell it at all, if you told a lie?

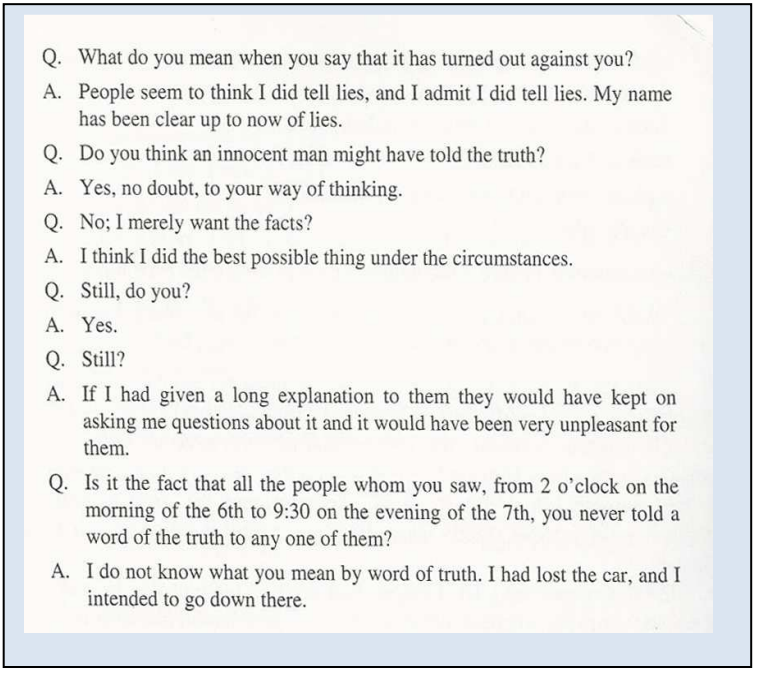
A. I was asked where my car was.

Q. You think it was unfortunate that you should tell lies in Wales?

A. It turned out to be subsequently, now, perhaps against me.

(a)

Figure 7(b): (continuation)



Q. What do you mean when you say that it has turned out against you?

A. People seem to think I did tell lies, and I admit I did tell lies. My name has been clear up to now of lies.

Q. Do you think an innocent man might have told the truth?

A. Yes, no doubt, to your way of thinking.

Q. No; I merely want the facts?

A. I think I did the best possible thing under the circumstances.

Q. Still, do you?

A. Yes.

Q. Still?

A. If I had given a long explanation to them they would have kept on asking me questions about it and it would have been very unpleasant for them.

Q. Is it the fact that all the people whom you saw, from 2 o'clock on the morning of the 6th to 9:30 on the evening of the 7th, you never told a word of the truth to any one of them?

A. I do not know what you mean by word of truth. I had lost the car, and I intended to go down there.

(b)

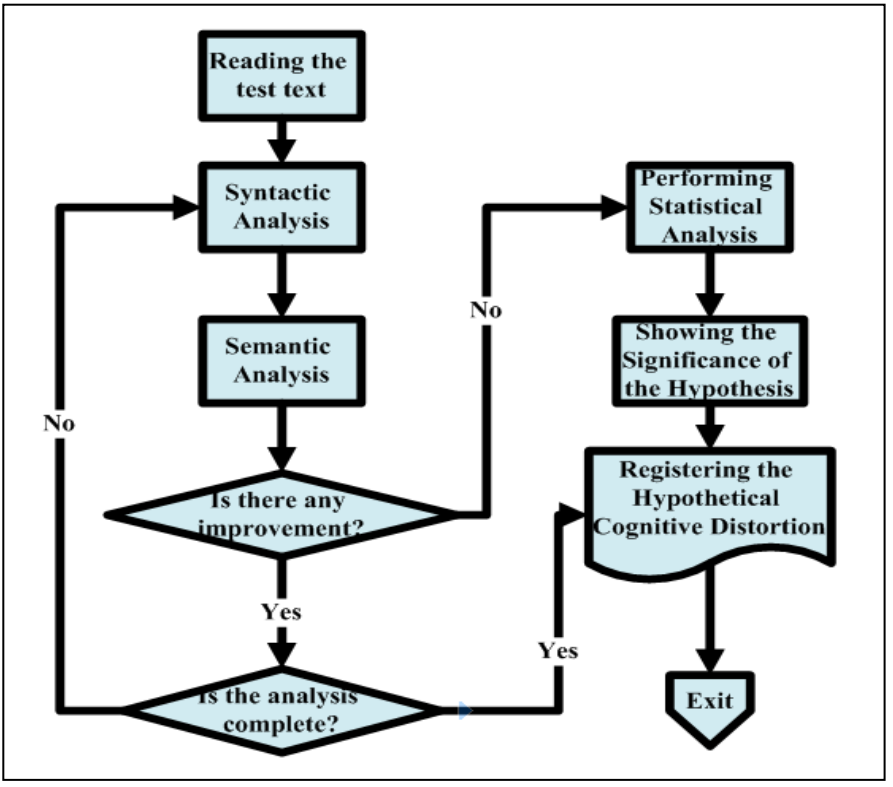
Evaluation: Identifying cognitive distortion

The *Quantitative-Semantics* (QS) and the BNF notation defined before enables the analysis of the sentence's meaning. Such analysis is essential for identifying thoughts having cognitive distortion. The structure (BNF) and the evaluation of meaning (QS) reinforce each other. Namely, BNF enables a more accurate way of classifying the sentence's words into their corresponding QS categories (Figure 8). In addition, conversely, the first-iteration of the sentence's word classification may improve the first decomposition.

Statistical tools, whose usage is demonstrated by describing the treatment of mental filter cognitive distortions, support the above analysis.

The next relevant step in *identifying cognitive distortion thoughts* (Apostolico and Galil (Editors) 1997, Navarr 2002, Charras 2004) is to discover whether the analyzed sentence belongs to one of the known categories in the bibliography (Burns 1999) of *cognitive distortion thoughts.*

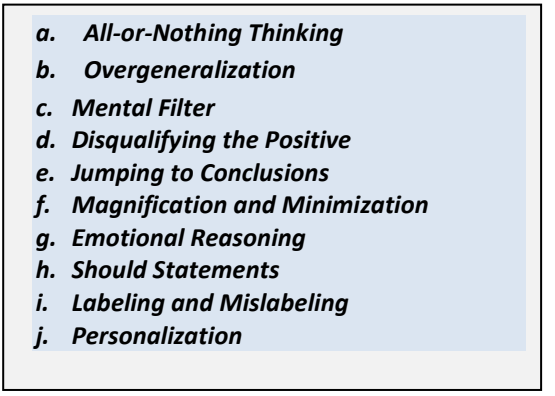
Figure 8: An algorithm for cognitive distortion recognition.



The thought distortions are generally recognized by using the *external-terms* (superlatives) introduced in section 3 in a special context, or by constructing a special sentence.

Some of the known distortions (Figure 9) will be listed below and their linguistic (semantic and syntactical) properties will be noted.

Figure 9: Cognitive distortion thought categories (Burns 1999).



(i) All-or-nothing thinking

A template (1) may determine this type of distortion.

(1) ***~ <maximal-term> , [<cause-connection>]***
I am(<minimal-still-term> | <negative-label-affront>)

The following example shows the essence of the above template (1.): "I received a grade of 85 in the examination. I am a complete fool; how could I make such a mistake?"

(ii) Overgeneralization

This type of distortion may be determined by a template (2) using the <timing-term> notations taken from Figure 1 and with the <event> notation, which denotes a sentence describing some kind of event.

(2) ***<overgeneralization-distorted-sentence> ::= <maximal-timing-term><negative-event>*** | ***<minimal-timing-term> <positive-event>***

The following examples use the ***overgeneralization*** class of thoughts:
"I ***always fail*** the examination."
"I ***never succeed*** in passing the examination."

(iii) Mental filter

This ***cognitive-distortion*** causes the person to perform a so-called ***selected-abstraction***. Here the template would be (3):

(3) ***[because] <negative-event> , <negative-relation>***

Example: "He laughed at her; that person is very cruel."

In order to identify this distortion, the tone can use statistical methods [11] to find significant use of *negative-relation*s in comparison with *moderate-relations* from the same person. The rate given may be compared with the rate known in the person's milieu (community/population).

(iv) Disqualify the positive
Here, the template consists of two components (4), where a <*positive-event*> generally describes a sentence in which the event yields a positive outcome for the subject. The component <*less-important-term*> relates to neglecting (an antonym of emphasizing) term.

(4) **<positive-event><less-important-term>**

An example is the following conversation:
You are a very good student. You received a very good grade.
It is a coincidence. I am just lucky.

(v) Jumping to a conclusion:
Drastic decisions are made, owing to bad speculations about the future and an incorrect reading of people's thoughts.

(vi) Minimization – maximization
This class is characterized by using the <*extreme-term*>*s.*

(vii) Emotional Reasoning
The category of cognitive distortion, *Emotional Reasoning*, is treated according to the template given next (5):

(5) ***I am <negative-emotional-term> <cause-connector>***
I am <minimal-still-term>

This definition can be improved by finding a more general determination of subject I.

(viii) Should statements
To treat the distortions based on *should statements*, we will perform ***pattern matching*** (see (Apostolico and Galil (Editors) 1997; Navarr 2002; Charras 2004)). The template ***<sharp-conscientiousness***> is related to the ***subject*** (the ***syntactic-part*** of a sentence), where the ***subject*** is referred to in ***first-person*** such as I and me.

(ix) Labeled and mislabeled
Labeled and mislabeled *sentences* constitute a class of distorted sentences containing the ***<negative-label-affront>*** attached to an ***object* (*syntactic-part*** of a sentence). This object relates to the second **or *third-person* (*speech-part***).

The difference between the *Labeled-terms* and ***Mislabeled-terms*** lies in the degree of the reliability. The mislabeled-term group members indicate the existence of ***distorted thought***, whereas the members belonging to the labeled-terms are candidates for causing disturbances through cognitive distortions. It all depends on the context in which the terms are used. The tree-structured analysis is helpful in analyzing such a context.

(x) Personalization

The cognitive distortion category, ***Personalization***, is treated according to the developed iCBT methodology. Additional items in the ***Personalization*** list should be treated as the others. This list should include the following expressions: self-blaming, negligence, fault, and responsibility. This is achieved by defining the template that treats the ***Personalization*** type of cognitive distortions.

Evaluation: Identifying cognitive distortion thoughts

The summarized procedure mentioned previously is given herein. The ***Quantitative-Semantics*** (QS) and the BNF notation defined before enables the analysis of the sentence's meaning. Such an analysis is essential for identifying thoughts having cognitive distortion. The structure (BNF) and the evaluation of meaning (QS) reinforce each other. Namely, BNF enables a more accurate way of classifying the sentence's words into their corresponding QS categories (Figure 8). In addition, conversely, the first-iteration of the sentence's word classification may improve the first decomposition.

Statistical tools, whose usage is demonstrated by describing the treatment of Mental Filter cognitive distortion, support the above analysis.

The next relevant step in ***identifying cognitive distortion thoughts*** is to parse (Apostolico and Galil (Editors) 1997; Navarr 2002; Charras 2004)a given transcription of thoughts to determine whether the analyzed sentence belongs to one of the known categories in the bibliography [2] of ***cognitive distortion thoughts.***

The thought distortions are generally recognized by using the *external-terms* (superlatives) introduced in section 3 in a special context, or in a special sentence construction.

Some of the known distortions (Figure 9) will be listed below and their linguistic (semantic and syntactical) properties will be noted.

(i) All-or-Nothing Thinking

A template (1) may determine this type of distortion.

(1) ***~ <maximal-term> , [<cause-connection>] I am(<minimal-still-term> | <negative-label-affront>)***

The following example shows the essence of the above template (1.): "I received a grade of 85 in the examination. I am a complete fool; how could I make such a mistake?"

(ii) Overgeneralization

This type of distortion may be determined by a template (2) using the <timing-term> notations taken from Figure 1 and the <event> notation, which denotes a sentence describing some kind of event.

(2) ***<overgeneralization-distorted-sentence> ::= I <maximal-timing-term><negative-event> | I <minimal-timing-term> <positive-event>***

The following examples use the ***Overgeneralization*** class of thoughts:
"I ***always fail*** the examination."
"I ***never succeed*** in passing the examination."

(iii) Mental filter
This ***cognitive-distortion*** causes the person to perform a so-called ***selected-abstraction***. Here the template would be (3):

(3) ***[because] <negative-event> , <negative-relation>***

Example: "He laughed at her; that person is very cruel."

In order to identify this distortion, the tone can use statistical methods (Stockburger 1996) to find significant use of *negative-relation*s in comparison with *moderate-relations* from the same person. The rate given may be compared with the rate known in the person's milieu (community/population).

(iv) Disqualifying the positive
Here, the template consists of two components (**4**), where a <*positive-event*> generally describes a sentence in which the event yields a positive outcome for the subject. The component <*less-important-term*> relates to a neglecting (an antonym of emphasizing) term.

(4) ***<positive-event><less-important-term>***

An example is the following conversation:
- *You are a very good student. You have received a very good grade.*
- *It is a coincidence. I just have good luck.*

(v) Jumping to Conclusions
Drastic decisions are made, owing to bad speculations about the future and an incorrect reading of people's thoughts.

(vi) Minimization and magnification
This class is characterized by using the <*extreme-term*>*s.*

(vii) Emotional reasoning
The category of cognitive distortion, *Emotional Reasoning*, is treated according to the template given next (5):

(5) ***I am <negative-emotional-term> <cause-connector>***
I am <minimal-still-term>

This definition can be improved by finding a more general determination of the subject I.

(viii) ***"Should*** " statements

To treat the distortions based on *should statements* we will perform *pattern matching* (Apostolico and Galil (Editors) 1997; Navarr 2002; Charras 2004). The template **<*sharp-conscientiousness*>** is related to the ***subject*** (the ***syntactic-part*** of a sentence), where the *subject* is referred to in ***first-person*** such as I and me.

(ix) Labeled and mislabeled

Labeled and mislabeled sentences constitute a class of distorted sentences containing the <*negative-label-affront*> attached to an *object* (*syntactic-part* of a sentence). This object relates to the second or ***third-person* (*speech-part*).**

The difference between the *Labeled-terms* and *Mislabeled-terms* lies in the degree of the reliability. The mislabeled-term group members indicate the existence of ***distorted thought***, whereas the members belonging to the labeled-terms are candidates for causing disturbances through cognitive distortions. It depends on the context in which the terms are used. The tree-structured analysis is helpful in analyzing such a context.

(x) Personalization

The cognitive distortion category, ***Personalization***, is treated according to the developed iCBT methodology. Additional items in the *Personalization* list should be treated as the others. This list should include the following expressions: self-blaming, negligence, fault, and responsibility. This is achieved by defining the template that treats the ***Personalization*** type of cognitive distortions.

Evaluation: Correcting cognitive distortion thoughts

After discovering the ***distorted thoughts*** and recognizing them in the transcribed text, the psychotherapist may suggest some corrections (Burns 1999) in the original text.

This operation may be partially computerized using known algorithms in the field of the string/tree *pattern-matching* (Apostolico and Galil (Editors) 1997, Navarr 2002, Charras 2004, Gawne-Kelnar 2008).

The proposed computerized ***CBT*** is intended to give the intended individual an opportunity to achieve gradual self-correction by choosing the appropriate expression from a list of proposed ***moderate-expressions*** (Figure 1) and substituting it for the ***extreme expression***. The ***iCBT*** can automatically perform such substitutions and show the user the computer's solution.

Concluding remarks

The whole cycle of the iCBT is schematically described in Figure 8. It should be emphasized that the adjectives and adverbs may be categorized into two classes: superlative and mild.

(i) Superlative

The ***superlative*** class is a class in which the terms can be categorized very easily, suggesting some ***cognitive-distortion***. This class contains expressions such as ***impossible*** or ***never***.

(ii) Mild

The ***Mild*** class contains expressions that express some doubt. Statistically, they more accurately describe reality (Burns 1999) and they may be substituted for the superlative counterparts. This class contains expressions such as *improbable* or *seldom.*

The iCBT ***(Computerized CBT)*** (Ophir 2012) *is* a kind of ***bibliotherapy*** (Weld 2009) that uses *reading* as a therapeutic treatment method. The presented methodology, together with transformational grammar (Chomsky 1957) (supported by statistical methods), transforms an affirmative sentence into an interrogative one, upgrades the *reading* to an interactive collaboration between the software-system and the user-client.

The advantages of *iCBT* over ***bibliotherapy*** lie in *iCBT's* adaptiveness and therefore, it responds more accurately to the client. In the future, an improved human-computer relationship using audio devices enabling voice recognition instead of the textual input devices will be used. These types of devices can be termed *media user interface (MUI)* instead of the current *graphic user interface (GUI)*, and will also include audio and other media possibilities.

A further suggestion is that future SHRE developments should quantitatively compare the SHRE results with that of the polygraph. A simple test would be to organize a group of volunteers who would be asked questions by the polygraphs operator. The answers given by the subjects would then be transferred to the software reliability tester and the evaluations can be compared with the polygraphs conclusions. It would be interesting to see the correlations between the conclusions of the two concepts: polygraph versus the SHRE.

Bibliography

Apostolico, Alberto and Galil, Zvi (Eds.) 1997. *Pattern matching algorithms*. Oxford: Oxford University Press.

Burns, David D**.** 1999. *Feeling Good*. Avons Books, Harper-Collins Publishers.

Charras, Christian. 2004. *Handbook of exact string matching algorithms.* London: College Publications.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Colma, George. 1970. *Cross-Examination, A practical handbook*. CapeTown: Juta & Company, Limited.

Cotterill, Janet. 1968. *Language in the legal process.* New York: Palgrave.

Ferdinand de Saussure and Roy Harris. 1986. *Course in General Linguistics*. Chicago: Open Court Publishing Co.

Gawne-Kelnar, Gabrielle. 2008. *Bibliotherapy*: *Book Yourself In*.

Gibbons, John. 2003. *Forensic linguistics: an introduction to language in the justice system.* Maiden, MA: Blackwell Publications.

Glisan, James Lindsay. 1991. *Cross-examination, practice and procedure*; Sydney: Butherworths.

Greenberger Dennis and Christine Padesky.1995. *Clinician's guide to mind over mood*. New York, NY: Guilford Press.

Kearns, Kate. 2000. *Semantics (Modern Linguistic*s). New York, NY: St. Martin's Press.

Knuth, Donald E. 1964**.** Backus normal form vs. backus naur form*.* Communications of the ACM 7 (12): 735-736.

Navarr, Gonzalo, 2002. *Flexible pattern matching in strings*. Cambridge: Cambridge University Press.

Olson, John. 2004. *Forensic linguistics: an introduction to Language, crime and the law*. London, New York: Continuum.

Ophir, Dan, 2012. *Computerized Cognitive Behavioral Therapy (CCBT)***.** EABCT- European Association for Behavioral and Cognitive Therapy, August 29th, 2012, Geneva, Switzerland.

Salhany, Roger E. 2006. *Cross-Examination, the art of the advocate*. Canada: LexisNexis Butterworths, Inc.

Shuy, Roger W., 1966*. Language Crimes: the Use and Abuse of Language Evidence in the Courtroom*; Cambridge, Mass.: Blackwell.

Stockburger, David W. 1996. *Introductory statistics: Concepts, models, and applications.* Atomic Dog Publishing.

Weld, Allen Hayden. 2009. *Parsing book*: Containing rules of syntax and models for analyzing and transposing, together with selections of prose and poetry. Whitefish, Montana: Kessinger Publishing, LLC.