Neuro-Linguistic Programming as a Verbal Lie Detection Tool: A Case Study of Pakistani Digital Media Programs

Hafsa Qadir[1]*, Dr. Zafar Iqbal Bhatti[2], Dr. Ayesha Junaid[3]

[1]*Ph.D. Scholar School of English, Minhaj University, Lahore
[2]Professor, School of English, Minhaj University, Lahore, Pakistan
[3]Assistant Professor, Center of Language Development, FCCU Lahore, Pakistan

*Corresponding Author: Hafsa Qadir
*Ph.D. Scholar School of English, Minhaj University, Lahore

Abstract
This research sightsees the application of neurolinguistic programming techniques by Pakistani fact-finding journalists with the accused and spectators in digital media programs. Besides, it extracts trickery indicators along with the frequencies. It also answers how the model is operative for the neurolinguistic lie detection system. These objectives are accomplished through a mixed-methods research paradigm. The data are collected from secondary sources, in which the accused, spectators, and inspecting journalists' turn-taking utterances are nominated for linguistic examination. For the case study, 2 case studies are taken from discussions in accountability reporting through purposive and convenient sampling. The researchers took 2 male and 2 female accused's spoken data. After that, the AI-powered 'TurboScribe' tool and 'Google Documents' are used for audio-to-text conversion. Truth and lie statements were extracted from the answers to relevant, irrelevant, and leading questions and were matched with neurolinguistics techniques. In rule-based matching, the XML-Roberta model was used on synchronized sentences to extract phrases or markers for dishonesty and reality. The linguistic patterns and sentence complexity (POS and types of sentences) were excavated. In light of these facts, the reporters use neurolinguistic techniques to extract valid information from the accused or suspects and witnesses. Furthermore, lie statements frequently comprise compound, complex, and compound-complex sentences while truth statements are habitually simple sentences. Moreover, the Roberta-base model is 75% effective for neurolinguistic deception detection.

Key Terms: Neuro-Linguistic Programming (NLP); digital media programs; accountability reporting, deception detection.

I. Introduction

The discovery of facts depends on the investigators' ability to raise effective questions. For this motive, investigators employ numerous techniques, including neurolinguistic interviewing techniques, which are frequently denoted as neuro-linguistic programming (henceforth NLP). It is based on verbal and non-verbal cues (Tucker, 2019). It includes an effective interpersonal communication model that improves communication between two or more persons (Biswal & Prusty, 2011).

NLP is an approach that was presented by Richard Bandler and John Grinder in 1975. The purpose of this technique is to divulge basic human interactive patterns. These co-founders were inspired by the transformational generative grammar theory given by Noam Chomsky in 1957. They formed a stepwise progression to scrutinize an individual's language, the 'surface' and 'deep' meanings of these structures, and presented therapy techniques to deal with these patterns (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton & Cooper, 2003).

The motives for leading this study are: foremost, the researchers are interested in discovering the accused' falsehood because law enforcement agencies reach out to different scientists and researchers to help them improve the criminal justice system. Secondly, law enforcement agencies rely only on NLP techniques (Tucker, 2019) for lie detection as it is a safer method than the harsh methods used by some investigators (Math, 2011).

Humans display truth and lie behaviors in their daily communication (Sarzynska-Wawer, Pawlak, Szymanowska, Hanusz, & Wawer, 2023). It is a challenging task to discriminate truth from falsehood because these behavioral patterns vary in speaking (DePaulo, Lindsay, Malone, Muhlenbruck, Charlton & Cooper, 2003). Zuckerman, DePaulo, and Rosenthal, (1981) say that lying is more complex than telling the truth because brain scanning studies found that the prefrontal cortex which controls the brain's functioning works more actively when a person is lying than truth-telling (Abe, 2011). Even, the fact originates in cognition primarily when somebody deceits (Debey, De Houwer, & Verschuere, 2014).

This matter has become more concentrated in recent years. Researchers in psychology, linguistics, and computer science seek behavioral deception cues. DePaul et al. (2003) defined lying as intentionally trying to trick others. Mistakes are not measured as lies, but statements that are exactly true but intended to betray are lies. Some researchers distinguish between deceiving and lying, but these are identical terms (DePaul et al., 2003).

The theme of truth and lies is often covered in digital media programs in Pakistan, particularly in political, crime, and social issues. These programs' content and shows are broadcast on different platforms, such as television, radio, and social media. It contains news, current affairs, entertainment, documentaries, morning shows, and talk shows that broadcast debates,

discussions, and reporting on events. The specialized reporters cover critical issues using neurolinguistic techniques to effectively present the truth to the public.

II. Literature Review

This research spins around Pakistani digital media programs in which investigation journalists use neurolinguistic programming to disgorge truth and detect lies from the accused or suspects, and victims' behavior. Crime is problematic worldwide. For this purpose, individuals are concerned with the pointers of convicts' deception in various fields through physiology, para-verbal, non-verbal, and verbal behaviors (Hauch, Gitlin, Masip, & Sporer, 2015). In Pakistan, wrongdoings are growing gradually. So, the accused or suspects' verbal statements are a unique source from which to study deception. Former studies associated with verbal lie detection cues are mentioned below:

Sarzynska-Wawer et al. (2023) verified the variances between truthful and false statements based on complexity and emotional factors using special tools designed to measure these features. The features of sentence length, concreteness, and parts of speech of truth from lie statements were observed in English and also in Polish. For this study, Polish-speaking participants were recruited through social media and online advertisements. The 400 participants were between 18 and 60 with primary, secondary, and higher-level education, of which 226 were females. With ethical approval, participants completed a questionnaire on 12 different topics, using a Likert scale in the first step. After that, participants generated two written and two oral statements by focusing on one topic: one is based on truth and the other on opposite beliefs. Approximately 1,600 true and false statements were analyzed with equal division. Oral statements were annotated automatically, and written statements were saved separately by remaining in the same condition. The findings showed that there is no vast variance between oral and written lies. The false statements are simpler, briefer, and more specific, had positive words, and rarer negative words. The results of the truth-and-false statements detection system received 60% correctness. It also contributes to former research on how lying is specified through language.

Loconte et al. (2023) discussed spoken lie detection utilizing large language models (henceforth LLM), also known as FLAN-T5 (standard and small size). This study detects lies from 3 English data records i.e., autobiographical, personal opinions, and future goals' memories. After analyzing the language differences in three English datasets, the researchers tested small and standard-size FLAN-T5 in three ways by using 10-way cross-validation, in which training and testing are done on the same single dataset. Secondly, training was completed on two datasets and tested on the third. Thirdly, training and testing were performed on all three datasets together. The best results were achieved in the first and third situations. The outcomes presented that the model's performance is enhanced with a larger size. Moreover, the stylometric analysis was performed for linguistic differences in the datasets. Hence the model made certain predictions. The results showed that the model's predictions are affected because language features are connected to the cognitive load framework.

Chung, Wong, Kamaluddin, and Bull (2023) investigated the interviewing techniques, i.e., relationship and rapport building, presentation of evidence and collaboration, emotional provocation, competition or confrontation, and context manipulation, with suspects by Malaysian police officers. A pen-and-paper survey was conducted on 88 police officers. The findings show that 70 percent of police officers interviewed with respect, patience, and kindness in building rapport, while 10% responded that they touched the interviewees in a friendly way. Additionally, police officers reported using methods like showing concern and appealing to the interviewee's desire for cooperation to build a sense of teamwork. More than half of the officers asked questions more than once from the interviewees using confrontational techniques. 75% of officers use minimization techniques, such as moral rationalization for interviewees' cooperation.

Vrij and Hartwig (2021) discussed the effect of wearing face masks on lie detection in the courtroom. During the pandemic, virtual courts were planned as compared to in-person court proceedings, and the jurors' task was to determine whether the defendants were guilty or not through non-verbal behavior. Due to medical face masks, non-verbal behavior is not possible to evaluate the truthfulness of the witnesses and defendants. In this situation, jurors pay attention to speech content to spot deception. The result concluded that virtual courts had no negative effect on jurors' lie detection abilities. Moreover, verbal cues are not affected due to the face mask. Still, face masks may suppress the defendants' speech, making it tougher for jurors to understand the speech and detect the lies. Verbal cues are reliable signs of deception, but only when particular interview methods are used that strengthen these cues, i.e., cognitive credibility assessment, verifiability approach, and strategic use of evidence.

Ask, Calderon and Giolla (2020) discussed human lie detection performance by comparing self-selected truth-tellers or liars versus those instructed to tell the truth or lie on a random basis. This research explored whether self-selected versus assigned progress weakens an individual's ability to detect deception. Experts criticized this approach because it does not reflect real-life situations. Both human performances (self-selected versus instructed) alter in various ways, i.e., proficiency of deception and motivation. These two ways can lead to different signs of truth and deception. 72 participants were selected for this research. The researchers discovered that when people are instructed to speak a lie, then it has a slight effect on human lie-detection performance and how accurately other people can spot their deception. When these results apply in real-life situations where individuals speak lies naturally, then it is still doubtful. It concluded that there is a difference in mental processing between natural or self-speaking lying and instructed lying.

Tucker (2018) analyzed the NLP methods (theories, usage, accuracy, evidence, problem, criticism, factors, and condition) from the available research to find their application and effectiveness in the criminal justice field. The diverse interview and interrogation methods were reviewed to know the theories behind NLP interviewing and its accuracy and to find the answers to the question of whether NLP can assist in crime effectively or not. It showed that investigation officers rely solely on neuro-linguistic interview techniques in finding the suspects by understanding verbal and non-verbal behaviors. However, there is limited research done on using NLP methods just to verify the guiltiness or innocence of a suspicion based on non-verbal behaviors (eye movement). After studying several methodologies of NLP interview techniques, it was decided that each of the

NLP techniques had the same results. Moreover, the NLP techniques have improved lie detection slightly more than chance in police departments through non-verbal behavior. Finally, it was suggested that a thorough analysis of criminals' deceitful behavior and the environment could be formed with the use of anthropologists because they study the environment to understand what influences and creates behavior change. So, investigation officers predict deceitful behavior in the interrogation rooms.

Hauch et al. (2015) investigated linguistic signs of deception by doing a meta-analysis using computers as an effective lie detector. The researchers united 79 cues from 44 studies to identify deception by using software. After that, these 79 cues were assembled into six research questions. Based on theoretical frameworks, the research used cognitive and social psychology theories to create and test particular ideas that might help in detecting deception by using linguistic cues. The results were significant in terms of how computers can help spot lies. The meta-analysis shows that deceptive people think harder, display more negative feelings, detach themselves from events, speak less about their own opinions, and use fewer descriptive or sensory-perceptual words. However, truthful people are more unclear than deceptive. These effects are moderated depending on the type of incident, participation, emotional tone, intensity of interaction, inspiration, and other factors. Although the complete effect size was small, certain theory-driven predicted cues were supported. These results support us in understanding how language can be used to detect lies by using computers and explain the construction between language and lying.

DePaulo et al. (2003) identified deception cues quantitatively. The main objective of the study was to examine the variance among truth-tellers versus liars. The researcher took 158 cues of deception with 1300 examples from 120 studies. The findings showed the following cues: first, deceptive people share less information, are less convincing, have strange details of the stories, and have fewer common faults. Secondly, liars have more negative feelings than truthful individuals. Thirdly, in a few situations, individuals' verbal behavior does not match or shows weak connections with deception. Fourthly, deceptive cues were prominent in those people who wanted success related to identity rather than money or things. Fifthly, when the lies are related to wrongdoings then the deception cues are more prominent in that situation. Furthermore, liars express themselves in more unclear and indirect ways than truth-tellers. Liars might display guiltiness and nervousness by distancing themselves from their lies. This concept was created by Wiener and Mehrabian (1968; 1972) about verbal and nonverbal signs of distancing. Liars take a longer time to respond and hesitate more when they are speaking. Moreover, their behavior seems less natural. There is a mismatch between words and actions in liars.

Sandoval and Adams (2001) explained that detectives and other experienced investigators know the critical importance of building rapport. During an investigation, rapport building and better communication are the most useful parts of NLP. NLP training in criminal justice is meant to help officers connect better with witnesses. The capability to connect well and build bonds is a key factor in the success of police officers when interacting with the community. People use language to convey their ideas or thoughts, and the selection of words tells us the way of their thinking. In the case of sharing experiences, people use sound, feeling, and sight to recognize and explain things to others. So, successful investigators listen to the words carefully. Officers are trained in controlling non-verbal and verbal patterns, as NLP can notice witnesses' body and speech patterns. Investigators match a person's speech patterns with paralanguage (including the rate, volume, and pitch). In an investigation setting, a detective slows down his speech to match a slow-speaking witness's, which helps him remember and share information better. When a witness's voice is quicker and louder, the investigator should try to match their energetic and expressive way of talking. An excited witness speeds up her speech rate and volume. As a result, it increases the chance of gaining relevant information.

Vrij and Lochun (1997) discussed that several books and studies talked about NLP, but nowadays, it is even used by the police. The study answered the questions about NLP, a literature review on NLP models, and its usefulness. NLP gives much information and reveals the PRS through non-verbal behavior (eye movement). NLP models the prerogative that individuals approach their brains through three senses: vision, kinaesthetic, and listening. One method that helps to know how people understand things is by listening to their words. Some people prefer the visual way of understanding, such as 'It seems like', while others choose kinesthetics, such as 'It feels like', some select the hearing modality, such as 'I listened how', and some might use all these three senses to understand the world. The weak point of NLP is that the co-founders of NLP, Bandler, and Grinder, did not research practically to check whether NLP is correct rather than focused on whether their ideas worked or not. The questions were asked as a point of departure for the participants to measure the verbal behaviors of the individual's truthfulness, such as "Do you have blue eyes?" So far, studies have focused on three modalities: sight, listening, and touch, but no one has tested whether this idea is correct.

Research Gap

This research sheds light on NLP techniques used by media journalists and their influence in disgorging truth and falsehood recognition. There are constrained studies in the criminal justice system about what people think and how they talk, compared to the huge amount of research about how people behave without talking. However, the investigators provision that it works through oral behavior for criminals. This study will analyze deceptive linguistic patterns used by the accused and suspects that will aid in revealing the truth in front of the public more easily.

Some researchers studied criminals' behavior through NLP (Tucker, 2019; Vrij & Lochun, 1997). These researchers recommended that linguistic patterns could also help in knowing criminal behavior. Previously, scientists have observed verbal cues that portray lies and generate ways to express truthful stories from deception. But these signs are minor, and even specialists only do slightly better than guess (DePaulo et.al., 2003; Loconte et al., 2023). Individuals are not good at detecting lies because they trust that others are truthful (Levine, 2014; Street & Masip, 2015). Finally, this type of research is not done in the Pakistani context. This study wants to fill this gap by studying the use of NLP techniques in Pakistani electronic media programs. It revolves around the accused or suspects' linguistic data that will help identify deceitful and truthful behavior. So, the above studies gave a clear future research potential and gap to this research.

III. Research Methodology

This section defines the population, sample size, sampling technique, data type (primary or secondary), data collection methods such as verbal or text, and research procedure. The qualitative and quantitative research paradigms are fitted in this research because spoken data was obtained from the criminals and accused or suspects from Pakistani media programs. So, it falls into the qualitative research paradigm. After that, the data was analyzed quantitatively. In short, the study design is mixed-method research. It mitigates the refraction of the researchers' subjectivity as NLP is a subjective approach labeled as quasi-science, which means it appears to be scientific but does not give any evidence. So, it avoids researchers' biases.

This research is chiefly exploratory, a type of qualitative research. It is a methodological approach that explores questions that have not been studied in detail. Verbal lie detection in the Pakistani criminal justice system is at a primary stage, aiming to understand speech patterns and find pointers of deception. In nature of data, the verbal data is collected from secondary sources through digital media programs, especially YouTube in the form of interviews, and discussions for linguistic analysis. The study's population includes investigative journalists, criminals, and suspects during media interviews. The population is extracted from media programs with 2 case studies. The data was collected through purposive or judgmental sampling because the researchers intentionally selected media programs related to research objectives. There are several programs related to crime, such as *Sar-e-Aam, Crime Patrol Pakistan, Shabbir to Dekhega, Andher Nagri, Juram Kahani, Aik Aur Jurm, Qatil Kaun,* and *Pukaar*. The researchers used the '*Pukaar'* media program because the reporter used a variety of neuro-linguistic techniques that fulfill the purpose of this research.

The researchers followed this procedure from data collection to analysis. The data were collected from the '*Pukaar'* media programs focusing on neurolinguistic techniques. Episodes were accessed via social media platforms i.e., YouTube. Secondly, the videos were downloaded using an online website *y2mate.com*. This website converted the video into audio form in MP3 format. Thirdly, the audio files were converted into text using the AI-based tool *TurboScribe.ai.* and *'Google Documents'*. The received data were in unstructured form. After that, the researchers pasted the data from Notepad to Excel format for manually cleaning the data. The researchers tagged the dialogues according to the utterances as a reporter, suspect I, suspect II, criminal, and police officer. Tags were fixed for ambiguities. In case of overlapping, it was tagged as 'unclear' in parathesis. Punctuation marks were added to mark the sentence boundaries in the turn-taking process where punctuation was missing. Meanwhile, borrowed words were also typed in the English language within parathesis.

In the data structuring process, the data was categorized into two parts: dialogues with neurolinguistics and without neurolinguistic strategies as conversations with suspects, criminals, and witnesses were with neurolinguistic techniques. The conversation of the investigation officers is conducted without neurolinguistic techniques. Additionally, the researchers separated the truth and lie sentences manually based on the investigative officer's statements, the first investigation report (henceforth FIR), and other shreds of evidence obtained by the media reporter from the data with the neurolinguistic technique because the data without neurolinguistic techniques are the facts of the story on the bases of which truth and lie statements are identified. The purpose of manual identification of truth and lie statements is that no trained or preexisting model is available for Pakistani languages. then the researchers used the *XML RoBERTa multilingual model* to detect truth and lie patterns from Urdu text data. In this process, rule-based lie and truth markers and keywords were extracted from the *RoBERTa* model from the English language. Subsequently, these keywords were translated into Urdu and integrated into the detection process to enhance contextual relevance. The purpose of this process is to check the accuracy of the model. The list of truth and lie phrases is mentioned in Table 3.2.

Table 3.2: Truth and Lie Phrases with Associated Markers

Truth Phrases	Lie Phrases
"بے شک", "یقینی", "واضح طور پر", "بالکل", "یقیناً", "یہ حقیقت ہے کہ", "یہ ایسا ہی ہوا تھا", "مجھے یقین ہے", "ایمانداری", "میں سچ کہہ رہا ہوں", "یہ بالکل ویسا ہی تھا", "تصدیق شدہ", "سچ", "حقیقی", "یہ حقیقتی صورت حال ہے", "مجھے حقیقت معلوم ہے", "میں تصدیق کر سکتا ہوں", "میں ہمیشہ تمہارے ساتھ ایماندار رہا ", "یہ بالکل واضح ہے ہوں"	"مجھے یقین نہیں ہے", "ہمم", "اِر", "آہ", "ارے", "شاید", "مجھے نہیں معلوم", "مجھے پختہ یقین نہیں ہے", "کہنا مشکل ہے", "ممکنہ طور پر", "یہ ہو سکتا ہے", "یہ سچ نہیں ہے", "کبھی نہیں", "غلط", "جھوٹ", "انکار", "مجھے یقین نہیں آ رہا", "میں نے یہ نہیں کیا", "یہ جھوٹ ہے", "مجھے یاد نہیں آ رہا", "میں یقین سے نہیں کہہ سکتا", "مجھے یاد نہیں کہ کہاں تھا", "یہ دھوکہ ہے", "یہ پیچیدہ ہے"

The researchers extracted 258 words for lie detection and 376 words of truthiness from the data. Later, neurolinguistic techniques used by the fact-finder were mentioned in front of truth and lie statements on the same Excel sheet. The files were separated for analysis as shown in Table 3.1.

Table 3.1 Truth and Lie Statements Analyzed Through NLP Techniques

Utterances	Turn-Taking	Statements	Neurolinguistic Technique	Status
ایف آئی آر (FIR) میں تو لکھا ہوا ہے کہ اپنے چچا سسر کے ساتھ تمہارے تعلقات تھے جس کی بنیاد پر تم نے اپنا شوہر قتل کروایا ہے.	Reporter			
یہ سب جھوٹ ہے. میں نے خود اپنے شوق سے کروایا ہے- میں نے اس کے ساتھ الزام لگا کہ ایسے کرنا تھا. یہ سب جھوٹ ہے.	Suspect	False	Yes (Pattern Interrupt)	Matched
یہی تو میں پوچھ رہی ہوں ہوں تم سے؟	Reporter			

Text	Speaker	Prediction		Match
نہیں آپی. سب جھوٹ ہے.	Suspect	False	yes (Meta-Model Questioning)	Matched
تو سچ کیا ہے؟	Reporter			
سچ یہی ہے کہ اس کو اس کے چاچو نے مار دیا تو مجھے نہیں معلوم کیوں مارا تھا کیوں نہیں۔ جگہ۔ جگہ کا.	Suspect	False	Yes (Pattern Interrupt)	Matched
ایک بیوی جو ہوتی ہے اس کی نظر عقاب کی طرح ہوتی ہے.	Reporter			
جی.	Suspect		yes (Presupposition Analysis)	Unmatched
اسے پتہ ہوتا ہے میرے شوہر کے حق میں اس کا کونسا رشتہ ٹھیک ہے اور کونسا نہیں ٹھیک۔	Reporter			
جی.	Suspect		Yes (Pattern Interrupt)	Unmatched
تو تمہیں پتہ ہی نہیں چلا. دو مہینے میں؟	Reporter			
تو میں اتنا اس کو. گونگا بہرہ تھا تو میں کیا اس کو اتنا سمجھتی تھی.	Suspect	False	yes (Meta questioning)	Matched
گھر والوں کو تو سمجھتی تھی نا؟	Reporter			
تو گھر والے میرے ساتھ ٹھیک تھے. وہ بھی ٹھیک ہے میرے ساتھ.	Suspect	True	Yes (Pattern Interrupt)	Matched

Part-of-speech (henceforth POS) tagging was executed using the Stanza Python natural language processing library to identify linguistic features and analyze grammatical structures associated with truthful and deceptive statements. For grammatical analysis, text was categorized according to the types of sentences based on structure, such as simple, compound, complex, and compound-complex. This process helped the researchers to evaluate the connection between sentence complexity and truthfulness or deception. This linguistic and sentence complexity analysis and extraction of indicators of trickery and reality markers contributed to improving the accuracy of detecting truth and lies in Urdu audio to textual data as shown in Figure 3.3.

Text	Speaker	Prediction	Sentence Type	POS Tags
دہ ماہ ہو گئے ہیں میں نے نے کروائی ہے اس کے ساتھ شادی.	Suspect		Simple	دہ/ADJ ماہ/NOUN ہو/VERB گئے/AUX ہیں/AUX میں/PRON نے/AUX نے/PRON کروائ
تمہارا یہ موجودہ شوہر ٹسیل (Disable) ہے؟ بول نہیں سکتا تھا؟	Reportor		Simple	تمہارا/PRON یہ/DET موجودہ/ADJ شوہر/NOUN ٹسیل/(Disable/PROPN)/
جی.	Suspect	Truth	Simple	جی/PUNCT.
سن نہیں سکتا تھا؟	Reportor		Simple	سن/VERB نہیں/PART سکتا/AUX تھا/AUX ؟/PUNCT
نہیں.	Suspect		Simple	نہیں/PUNCT.
گھر والوں کو تو سمجھتی تھی نا؟	Reportor		Simple	گھر/NOUN والوں/ADJ کو/ADP تو/PART سمجھتی/VERB تھی/AUX نا/NCT
تو گھر والے میرے ساتھ ٹھیک تھے. وہ بھی ٹھیک ہے میرے ساتھ.	Suspect	Truth	Simple	تو/SCONJ گھر/NOUN والے/ADJ میرے/PRON ساتھ/ADP ٹھیک/ADJ تھے/ X/.
تب تمہیں اس کی سمجھ آتی تھی اشاروں سے؟	Reportor		Simple	تب/PRON تمہیں/PRON اس/PRON کی/ADP سمجھ/NOUN آتی/VERB تھی/X
جی-	Suspect	Truth	Simple	جی/-PUNCT
ان کے معاملات کیا؟ جب تمہاری شادی ہوگی تو تمہارا بھی معاملہ تھا؟ تو بتاو ک	Reportor		Compound	ان/PRON کے/ADP معاملات/NOUN کیا/VERB ؟/PUNCT جب/PRON تمہارا
توکبہ رہی ہوں ناکہ زمین کا معاملہ.	Suspect	Truth	Complex	توکبہ/PROPN رہی/VERB ہوں/AUX نا/SCONJ کہ/SCONJ زمین/NOUN کا/؟
کیا معاملہ تھا زمین کا	Reportor		Compound	کیا/DET معاملہ/NOUN تھا/AUX زمین/NOUN کا/ADP
ان کے چاچو نے مارا ہے. ان سے پوچھیں۔	Suspect		Simple	ان/PRON کے/ADP چاچو/PROPN نے/ADP مارا/VERB ہے/AUX -/PUNCT
یہ دکھ سے آتی ہوئی ہو؟ تمہارے خاوند کو ایک بندہ مار دیتا ہے اور تمہیں پتہ ہو	Reportor		Compound	یہ/DET دکھ/NOUN سے/ADP آتی/VERB ہوئی/AUX ہو/VERB ؟/PUNCT تم
رات کو نہیں وہ میرے ساتھ تھوڑا سا بولتا تھا اس نے اپنا موبائل توڑا تھا تو وہ	Suspect		Simple	رات/NOUN کو/ADP نہیں/PART وہ/PRON میرے/PRON ساتھ/ADP توڑا/ET/

Figure 3.3 Visual Representation of Linguistic Analysis

After that, the questions were detected as relevant, irrelevant, and leading questions with frequencies because they help in the analysis of psychology, communication, and behavior and help to identify the complexity of sentences (simple, compound, complex, and compound-complex) in leading, relevant and irrelevant questions. It was identified by giving rules as shown in Figure 3.4

```python
# Function for NLP-based behavioral analysis
def detect_patterns(text):
    doc = nlp(text)
    matches = matcher(doc)
    detected_labels = set()
    for match_id, start, end in matches:
        label = nlp.vocab.strings[match_id]
        detected_labels.add(label)
    return detected_labels

# Define enhanced keywords for question type detection
question_patterns = {
    "IRRELEVANT": ["نام", "background", "state of mind", "routine"],
    "RELEVANT": ["عمر", "circumstances", "specific details", "crime"],
    "LEADING": ["کیا", "yes", "no", "confirmation"],
}

# Function to detect question type
def detect_question_type(question):
    question = question.lower()
    for q_type, keywords in question_patterns.items():
        if any(keyword in question for keyword in keywords):
            return q_type
```

Figure 3.4 Visual Representation of Type of Questions

Finally, NLP techniques such as pattern interrupt, cause and effect analysis, meta-modeling, rapport building, and sensory-based language analysis were applied to detect subtle linguistic and psychological cues in truthful and deceptive statements to enhance the accuracy of the detection process in the Urdu language. The procedure is described in the flowchart in Figure 3.4
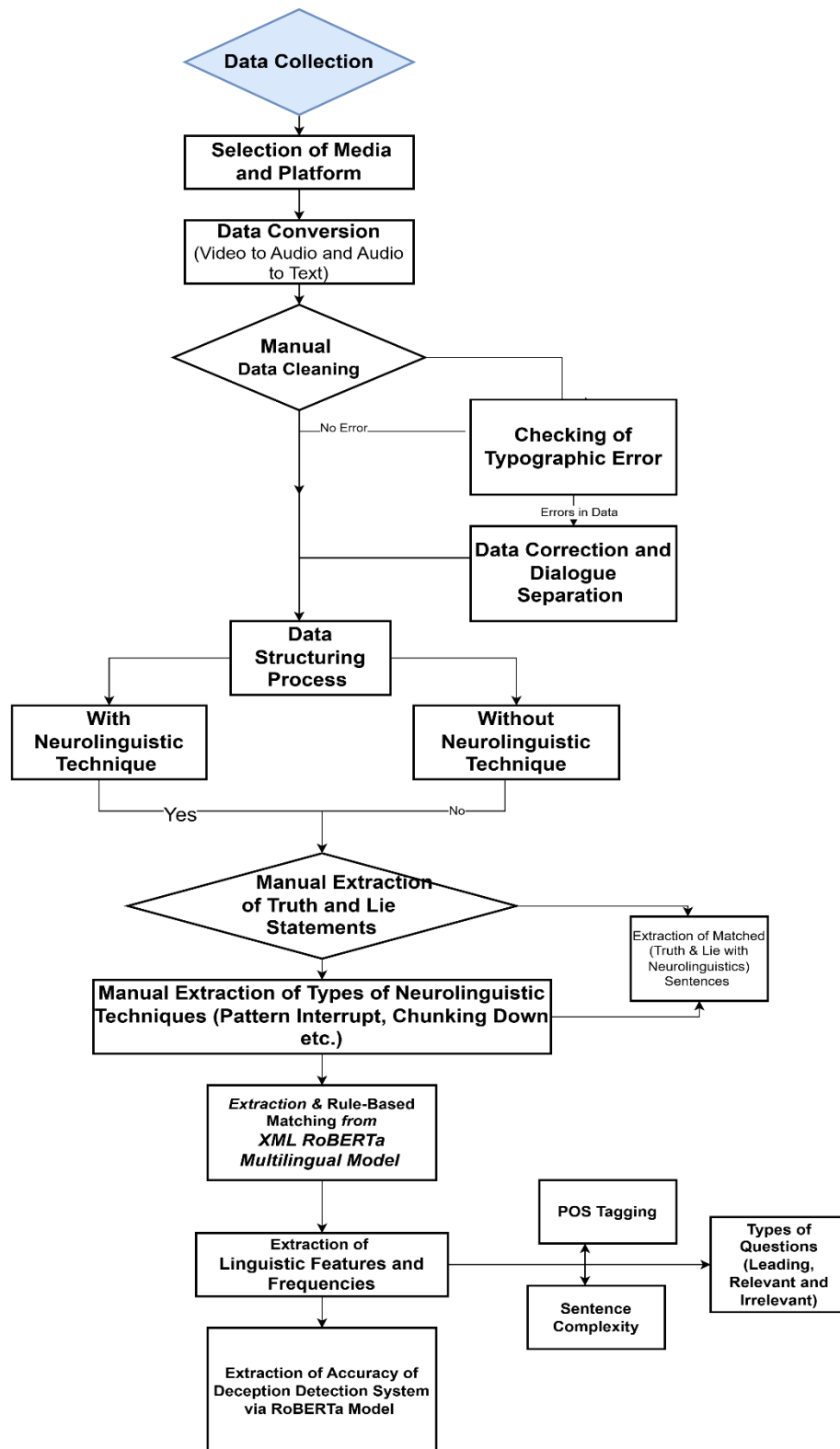

Figure 3.3 Visual Representation of Data Collection and Analysis Process

IV. Discussion and Results

Neurolinguistics research is at the cutting edge of scientific development by connecting the fields of natural sciences and linguistics. Pakistan's investigative journalists are skilled in constructing questions to obtain clear and comprehensive information from suspects, spectators, and victims. Their effectiveness is considered when they can persuade respondents to reveal truthful information. Reporters employ several types of questioning methods, not only case-relevant and appropriate

questions but also introduce diverse questions to get a fuller portrait of the case. Such questions aim to uncover facts, identify inconsistencies in testimony, and seek the truth about the crime.

Table 4.1. Types of Questions

Sr. #	Types of Questions	
1.	**Irrelevant Questions**	Background and routine information, state of mind
2.	**Relevant Questions**	Circumstances of crime
3.	**Leading** or **Optional Questions**	Yes/No questions

In the first murder case, the investigative journalist orally delivered 307 utterances with neuro linguistic techniques out of which 168 questions were delivered to the criminal and 139 were delivered to the suspects. In the second case, 163 questions were asked out of which 80 questions were asked from the criminal, and 77 questions from witnesses and neighbors. The total utterances were 627.

The female reporter investigates with neutral, accusatory, and leading questions. While the criminals' responses were from defensive to uncertain. The relevant questions are those questions that directly connect with the case. Such as "کتنی عمر ہے؟" Irrelevant questions are related to background information, routine, and associated with the state of mind, such as " کیا نام ہے. While leading questions are yes/ no questions as "تو پھر کیا تکلیف ہو گئی تھی شوہر سے جو تم نے اتنا بڑا قدم اٹھا لیا؟ تمہارا؟".

The neurolinguistic techniques are: nominalizations, probing and elicitation, reassurance or minimization, vague or ambiguous language, embedded command or suggestion, mirroring, cause-and-effect relationships, emotional cues, and provocation, anchoring and association, guilt induction and negative reinforcement, implication, and presupposition, social proof, and implied norms, tag questions and deflection, third-party validation, leading questions vs open-ended, reframing, outcome-oriented focus, sensory-based language, pacing, and leading, dissociation and distancing, meta-model questioning, double bind, chunking down, mind reading, emotional probing, evasion, mirroring, challenging identity, blame frame, shifting responsibility, challenging beliefs, calibration, and elicitation, confrontation, curiosity and engagement, subtle influence, curiosity and implied judgment, simplicity and directness, distinction, minimization, authority, confirmation, future pacing, emotional states, confirmation-seeking, implicit expectation, assumed responsibility, assumption of shared knowledge, contextual, emotional provocation, layered questions, pattern interrupt technique, creating cognitive dissonance, empathy, and rapport building.

First, the pattern interrupt technique was used to break the usual pattern of thinking of the criminal. This technique makes criminals more open, and responsive. By using this technique, the respondent became aggressive and committed the crime as shown in Table 4.2

Table 4.2: Example of Pattern Interrupt Technique

Utterances	Turn-Taking
ایک نشائی کے ساتھ اس کی بیوی نہیں رہنا چاہتی تووہ کردار کی غلط ہے – جو آپ نشہ کرتے ہیں تو اس کی اخلاًقی یا شرائی طور پر اس کی اجازت ہے۔	**J.**
e:k nəʃa:i ke sa:tʰ ʊs ki: bi:vi: nəhĩ: ræhna: tʃa:hʈɪ̈: to: ʋo: kɪrda:r ki: ɣələʈ̪ hæ. dʒo: a:p nəʃa: kərte hæ̃: to: ʊs ki: ɪxla:qi: ja: ʃərəi: tɔ:r pər ʊs ki: ɪdʒa:zət hæ	**IPA**
کیا وہ کر لے یا کروا لے	**C.**
kɪja: ʋo: kər le: ja: kərʋa: le:	**IPA**
کیا اجازت ہے؟	**J.**
kɪja: ɪdʒa:zət hæ	**IPA**
نہیں۔	**C.**
nəhĩ:	**IPA**
اگر وہ اس وجہ سے طلاق لینا چاہتی تھی تو وہ غلط ہے؟	**J.**
əgər ʋo: ɪs ʋədʒa: se: t̪əla:q le:na: tʃa:hti: t̪ʰi: to: ʋo: ɣələʈ̪ hæ	**IPA**
میم آپ یہ سوال۔۔	**C.**
Mæm a:p je: səʋa:l….	**IPA**
سوال ہے میرا جواب دیں ؟	**J.**
səʋa:l hæ me:ra: dʒəʋa:b dẽ.	**IPA**
میں نے جو کرنا تھا میم وہ میں کر دیا اور میں نے اوپنلی کیا ہے۔ ٹھیک ہے نہ۔	**C.**
me:ne dʒo: kərna: θa: Mæm ʋo: mẽ kər d̪ɪ̈ja: ɔ:r me: ne: oʊpənli kɪja: hæ - t̪ʰi:k hæ nə:	**IPA**

Secondly, a meta-model of language neuro-linguistic technique was used in which criminals share generalized, deleted, and distorted (incomplete information). The reporter employed this technique to excerpt necessary and detailed information that the criminal was hiding as shown in Table 4.3.

Table 4.3: Example of Meta-Model Questioning

Utterances	Turn-Taking
میں کافی ٹائم سے گھر میں خالی بیٹھا ہوا تھا	C.
mæ̃ ka:fi: daɪr se: gʰər me: xa:li: baeʈʰa: hua: ʈʰa:. laɪk	IPA
کیوں، کام کیوں نہیں کر رہے تھے	J.
kɪõ:, ka:m kɪõ: nəhĩ: kər rəha: θe:	IPA
بس جی کچھ گھریلو ناچاکی کی وجہ سے ذہینی طور پر اپ سیٹ تھا	C.
bəs dʒi: kʊʧ gʰəre:lu: nɑ:ʧa:ki: kɪ ʋədʒə se: zəhe:ni: tɔ:r pər əp sæ:t ʈʰa:	IPA

The journalists used rapport-building techniques with the criminals. This technique helps the reporter to develop trust and extract information from the respondent. The reporter did the same as shown in Table 4.4

Table 4.4: Example of Rapport Building Technique

Utterances	Turn-Taking
ماں، باپ، بھائی یہ وہ رشتے ہوتے ہیں جنہوں نے بتانا ہوتا ہے کہ	C.
mã:, bɑ:p, bʰa:ɪ, jɪʌ ʋo: rɪʂte: hote: hĩ: dʒĩnʰo: ne: bəta:na: hota: hɛ: kɛ:	IPA
اچھا کیا ہے برا کیا ہے	J.
əʧʰɑ: kɪja: hæ: bʊrɑ: kɪja: hæ:	IPA

Chunking Down Technique
This technique uncovers relationships and roles and clarifies the incomplete or vague responses of the accused or suspects as shown in Table 4.5.

Table 4.5: Example of Chunking Down Technique

Utterances	Turn-Taking
کتنے بچے ہیں؟	**J.**
kɪʈne: bəʧʧe: hæ̃:	**IPA**
تین	**C.**
ʈi̪:n	**IPA**
جو قتل ہوا اس سے کیا رشتہ تھا؟	**J.**
dʒo: qəʈl hʊɑ: ʊs se: kɪjɑ: rɪʂhʈɑ: ʈʰɑ:	**IPA**
بھتیجہ ہے میرا.	**C.**
bʰəʈi̪:dʒɑ: hæ: me:rɑ:	**IPA**
سگا بھتیجہ ہے.	**J.**
səgɑ: bʰəʈi̪:dʒɑ: hæ:	**IPA**
ہاں جی.	**C.**
ɦɑ̃: dʒi:	**IPA**

Linguistic complexity and parts of speech analysis were conducted to identify indicators of trickery in suspects' or accused/criminals' responses. In truth neurolinguistic statements, 52 out of 90 were simple sentences in which leading questions were asked, while 8 were relevant and irrelevant, as shown in Table 4.6.

Table 4.6: Analysis of Truth Neurolinguistic Statements

Statements	Truth			
Type of Questions	Leading/ Questions	Tag	Relevant Questions	Irrelevant Questions
Type of Sentence (Based on Structure)	Simple		Simple	Simple
Total	52		8	5

Truth and lie patterns vary in speaking. The *RoBERTa* model analyzed the sentences based on word dependency. The false statement was not extracted by the *RoBERTa* model. So, the researchers took a false statement that was extracted manually for sentence complexity. In false neurolinguistic statements, 5 sentences out of 27 were simple, 4 compound and complex sentences while compound-complex had one sentence in which leading questions were asked, while relevant questions were simple, compound, and complex sentences with the distribution of 5, 1, and 4. Irrelevant has one simple sentence as shown in Table 4.7.

Table 4.7: Analysis of False Neurolinguistic Statements

Statements	False		
Type of Questions	Leading/ Tag Questions	Relevant Questions	Irrelevant Questions
Type of Sentence (Based on Structure)	Simple, compound, complex and compound-complex	Simple, compound, and complex	Simple
Total	14	12	1

The following are the indicators that are not found in the RoBERTa model and present in the case study of 2 female and 2 male criminals or suspects' data as shown in Table 4.8

Table 4.8: Indicators of Truth Statements with Frequencies:

Words with POS Tags	Statements	Frequencies
X/جی NOUN/ہاں	Truth	49
نہیں PART/نہیں PART	Truth	2
مجھے/PRON کیا/PRON پتا/VERB	Truth	2
جی/PROPN جی	Truth	2
مینوں/PRON ضرورت/NOUN کی/ADP ہے/AUX	Truth	1
مینوں/PRON ضرورت/NOUN کی/ADP اے	Truth	1
میں/ADP نے/PART نہیں ADP/ سوچا/VERB تھا/AUX	Truth	1
عقل/NOUN تے/ADP پردہ/	Truth	1
غلطی/NOUN ہو/VERB گئی/AUX ہے/	Truth	1
ایہو/PROPN تے/ADP غلطی/NOUN ہو/VERB گئی	Truth	1
توکہہ/PROPN رہی/VERB ہوں/AUX نا	Truth	1
نہیں PART/- PUNCT/استغفار/NOUN ۔	Truth	1

False statements include the following indicators as shown in Table 4.9

Table 4.9: Indicators of False Statements with Frequencies:

Words with POS Tags	Statements	Frequencies
مجھے/PRON نہیں NOUN/ پتا/-VERB	False	3
پتہ/NOUN نہیں PART/	False	2

Issues, Challenges and Solutions:

The researchers faced multiple challenges during data processing which are mentioned below:

1. When the audio files were converted into text using the AI-based tool *TurboScribe.ai.* and *'Google Documents'*. These transcription tools replaced with the incorrect word, such as پڑی، نہیں نہیں، گنگار، گونگا، نئی نئی as پڑھی، نئی نئی. Then the researchers manually correct the words.

2. Transcription tools converted most of the sentences into textual data but Punjabi words were not typed and sometimes, English language words were also not typed in the text form, such as 'behave'. The researchers corrected the mistakes manually.

3. When identifying the pattern interrupt technique through Python programming, the researcher faced the challenge that the pattern was not automatically recognized because textual data was recognized based on full stop or comma, where stop words are not available. The researcher recognized the words by listening to the audio and then processing the file.

4. In some situations, truth and lie exist in the same sentence then the researchers separated those sentences into separate utterances.

5. Some utterances were based on non-verbal and verbal communication but the researchers marked the utterances based on the reporter's remarks as mentioned below:

تمہاری کتنی عمر ہے؟
35-40 سال عمر ہے۔
اچھا یہ 35 سال کی خاتون ہمارے ساتھ موجود ہیں اس کے بعد میں ان کی طرف آتی ہوں۔ چلو آگے آجاو پتہ نہیں تم آگے بھی سچ بولوگی کہ نہیں۔ اچھا۔ مجھے یہ بتاؤ کہ تم نے معاونت کیوں کی۔

6. In textual data, if there are multiple questions in single utterances then it isn't easy to understand the response either respondent answer to the first question or second as mentioned below:

اپنی اولاد کو مار دو گے؟ تمہارا بیٹا ہے؟
ہاں جی میری بیٹے دو ہیگے ہیں۔

7. Sometimes, it is difficult to understand whether the sentence is an interrogative one or a declarative but punctuation marks were marked based on listening to the audio as

ابھی نئی اس کی بیوی بیعاح کے آئی تھی۔

باں جی۔ وہ شادی میں نے کی ہے۔ وییاح کے میوں لییاندی اے اس کو۔

8. The *RoBERTa model* is successful in identifying truth statements but not lies. Then, lie statements were recognized manually.

Recommendation

Future research could develop speech-to-text corpora for criminal investigations for law enforcement agencies and the judicial process in Pakistan. In the current criminal justice system, additional and session judges rely on manual contradictions in utterances between witnesses and suspects, which slows the decision-making process.

References

9. Abe, N. (2011). How the brain shapes deception: An integrated review of the literature. *Neuroscientist, 17,* 560–574. doi:10.1177/1073858410393359.

10. Ask, K., Calderon, S., Mac Giolla, E. (2020). Human lie-detection performance: Does random assignment versus self-selection of liars and truth-tellers matter? *Journal of Applied Research in Memory and Cognition*, *9*(1),128-136. https://doi.org/10.1016/j.jarma

11. c.2019.10.002

12. Biswal, R., & Prusty, B. (2011). Trends in neuro-linguistic programming (NLP): A critical review. *Social Science International, 27*(1), 41-56. Retrieved from https://www.proquest.com

13. Chung, K.L., Wong, Y. S., Kamaluddin, M. R., & Bull, R. (2023). Investigative interviewing with suspects: Exploring current practices among Malaysian police. *International Journal of Police Science & Management*, *26*(1), 118- 128. https://doi.org/10.1177/14613557231205785

14. Debey, E., De Houwer, J., & Verschuere, B. (2014). Lying relies on the truth. Cognition, 132, 324–334. doi:10.1016/j.cognition. 2014.04.009.

15. DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological Bulletin,129*(1),74-118. https://doi.org/10.10 37/0033-2909.129.1.74

16. Hauch, V., Blandón-Gitlin, I., Masip, J. & Sporer, S. L. (2015). Are computers effective lie detectors? A meta-analysis of linguistic cues to deception. *Personality and Social Psychology Review*, *19*, 307–342. https://doi.org/10.1177/1088868314 556539

17. Math, S. B. (2011). Supreme Court judgment on polygraph, narco-analysis & brainmapping: A boon or a bane. *The Indian Journal of Medical Research,134*(1), 4-7 https://cqu-primo.hosted.exlibrisgroup.com/permalink/f/1rb43gr/TN_cdi_pubmedcentral_primary_oai_pubmedcentral_nih_gov_3171915

18. Sarzynska-Wawer J, Pawlak A, Szymanowska J, Hanusz K, Wawer A (2023) Truth or lie: Exploring the language of deception. *PLOS ONE 18*(2): e0281179. https://doi.org/10.1371/journal.pone.0281179

19. Street, C. N. H. & Masip, J. (2015). The source of the truth bias: Heuristic processing. *Scandinavian Journal of Psychology, 56*(3), 254–263. https://doi.org/10.1111/sjop.12204.

20. Levine, T. R. (2014). Truth-default theory (TDT): A theory of human deception and deception detection. *Journal of Language and Social Psychology*, *33*, 378–392. https://doi.org/10.1177/02619 27x14 535916.

21. Loconte, R., Russo, R., Capuozzo, P., Pietrini, P., & Sartori, G. (2023). Verbal lie detection using Large Language Models. *Scientific Reports, 13*(1), 22849. Retrieved from https://www.researchgate.net/

22. Tucker, M. (2019). Understanding Neurolinguistics Interviewing Techniques among Law Enforcement Agencies. Unpublished manuscript

23. Vrij, A., & Hartwig, M. (2021). Deception and lie detection in the courtroom: The effect of defendants wearing medical face masks. *Journal of Applied Research in Memory and Cognition, 10*(3), 392–399. https://doi.org/10.1016/j.jarmac.2021.06.002

24. Vrij, A., Lochun SK (1997). Neuro-linguistic programming and the police: Worthwhile or not? *Journal of Police and Criminal Psychology, 12*(1), 25–31. https://doi.org/10.1007/BF02813808

25. Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 14, pp. 1–59). New York: Academic Press