Hesitations and Unfilled Pauses:

An Empirical Analysis of their Use in Deceptive and Non-deceptive Speech

Schriftliche Hausarbeit
für die Bachelorprüfung
der Fakultät für Philologie
an der Ruhr-Universität Bochum

(Gemeinsame Prüfungsordnung für das
Bachelor/Master-Studium im Rahmen des 2-Fach-
Modells an der RUB vom 16.09.2014)

25.10.2020

Prof. Dr. Christiane Meierkord

Dr. Verena Minow

Vorgelegt von:
Herrmann, Sabine

Matrikelnr.: [redacted]
[redacted]@rub.de

Table of Contents

1. List of Tables, Graphs and Figures

1.1 Tables

1.2 Graphs

1.3 Figures

2. Introduction

2.1 Introduction and Research Question

People lie every day (DePaulo et al. 1996: 979). The average person lies about one to two times a day, though most of lies tend to be "everyday lies" of little relevance, with more serious lies occurring on a far less regular basis (DePaulo et al. 1996: 979, 993). Naturally, the concept of deception detection becomes relevant for most people at numerous points throughout their lifetime. Whether it is a law enforcement officer tasked with interviewing a suspect, a businessman trying to strike a lucrative deal or a wife suspecting her spouse of infidelity, the ability to spot a lie is arguably crucial to ones own and other's personal, social and financial security. Because of this, researchers have been trying for decades to find patterns from which to determine an individual's honesty or dishonesty, yet they have failed to reach a unanimous conclusion. Whilst early research focused predominantly on physiological cues such as micro-expressions (Ekman and Friesen 1969) or gaze aversion (Keltner and Harker 1998), the importance of linguistic and paralinguistic cues has become more apparent over time. Recent evidence suggests that changes within an individual's speech pattern, focusing on pitch (Zuckerman et al. 1981), hesitations (Kasl and Mahl 1965) or content and word choice (Newman et al. 2003, Nicklaus and Stein 2020) could potentially point towards deception taking place. There is also indication that listeners perceive the use of hesitation markers as indicative of lack of knowledge (Brennan 1995: 395-397) or pauses without hesitation markers as a sign of increased anxiety (Christenfeld 1995: 182). As such factors are commonly associated with deception (Mann 2019: 410), one has to wonder if these linguistic cues truly are reflective of the speaker's state of mind.

In the aim to explore this question this paper will analyze the prevalence of uh/um hesitation markers and pauses within both truthful as well as deceptive statements made by infamous fraudster Elizabeth Holmes, former founder and CEO of Theranos in a number of recorded interviews. In doing so this paper will investigate for potential changes found within the two speech scenarios to explore the possibility of uh/um hesitation markers and/or pauses being reliable indicators for deception. In extension to that this paper aims to uncover more evidence for the role of uh/um hesitation markers in linguistic theory and language proper. More specifically, where are hesitation markers employed within sentences structure? Do they occur arbitrarily, or can one recognize a pattern of distribution? Could such a pattern provide insight into the lexical role of hesitations? Namely, are uh/um hesitation markers mere

symptoms of cognitive efforts or are they strategically employed by the speaker and thus have recognizable syntax? Furthermore, is there any recognizable relation between the placement of uh/um hesitation markers and silent pauses?

2.2 Structure

First, this paper will give a short introduction to Elizabeth Holmes and Theranos to give a better understanding as to why her statements were chosen as a source for the following analysis. The paper will then go on to shortly explain the significance of deception within linguistic theory as well as the role of pauses and uh/um hesitations in language and discourse. Here the paper will illustrate the discussion on the lexical properties (or lack thereof) of uh/um hesitation markers, which lead to the divide in opinions on the question of hesitation markers being either 1) symptoms of cognitive load or 2) interjections with consistent properties. Following, the paper will briefly summarize the research conducted on the role of uh/um hesitations and pauses within deceptive speech. In accordance with this previous research, the hypothesis and expected findings regarding possible results are presented. The next section then details the data and methodology this paper uses. Following this, the results of the analysis are presented. Lastly, these results will be discussed and their significance with previous research will be evaluated. Here suggestions for further research and improvements on the methodology within this paper are given as well.

3. Context

3.1 Words Lie, Blood doesn't: The Story of Theranos

To better understand the subject of analysis, this paper will give a short introduction on Elizabeth Holmes and her company Theranos and why this paper uses her as source material for the truthful and deceptive speech scenarios.

Elizabeth Holmes dropped out of Stanford in 2003 to start Theranos, a bio-tech company focusing on blood-analysis work. Theranos claimed to have developed a machine, called the Edison, which allowed for more accessible and cheaper blood-analysis that could be conducted at home. Traditionally, a patient must go see a physician who then draws several vials of venous blood from the arm of the patient to send off to a laboratory for further analysis, which is a costly process in the US. The Edison machine was claimed to run the same tests using only a small amount of blood from a finger-prick test. The machine would then automatically conduct the analysis and offer immediate results without the need of a physician or lab assistant being present, drastically reducing the price. Theranos wanted to implement an Edison machine into every American household, as well as their own "wellness-centers" across the country. Deemed highly valuable by investors the company soon rose to a peak value of $10 billion in 2013/14. Theranos' board was filled by prestigious members such as former Head of State George P. Schultz and Henry Kissinger, luring in more investors and gaining Holmes celebrity status in Silicon Valley, often describing her as the new Steve Jobs. During this time Holmes also gave numerous interviews some of which this paper will use as source material (see 6. Data and Methodology). What investors were unaware of however was that the promise of the Edison machine was physically impossible to fulfill. In short, blood-analysis conducted via a finger-prick test can only analyze for a selected range of health issues (such as HIV or glucose-levels) and is prone to inaccuracies due to the small sample size. Furthermore, the Edison machine itself was highly unreliable and often delivered faulty results, forcing Theranos to secretly conduct their analysis in their own in-house lab using traditional Siemens products, a fact that was hidden from the public. Theranos was essentially selling a product that would never go live whilst testing it on cancer patients, supplying them with false health information data. The fraud was eventually uncovered in late 2015 by whistleblower Tyler Shultz[1] and John Carreyrou[2], investigative

[1] Grandson of board member George P. Shultz
[2] John Carreyrou details his findings in his book *Bad Blood: Secrets and Lies in a Silicon Valley Start-up* (2018)

journalist for the Wall Street Journal. This investigation in turn lead up to Elizabeth Holmes being charged with massive fraud by the SEC and multiple counts of wire fraud by the Department of Justice in 2018.

3.2 Locating Deception: Why Holmes?

Holmes was chosen as a subject for the analysis for several reasons. First of all, when planning for this paper it became soon apparent that there is a simple lack of reliable corpus data in regards to truth-telling versus lie-telling. Of the very few corpora that exist majority deal with a highly specific set of data. The Boulder Lies and Truth Corpus (BLT-C) for example concerns itself solely with written text in the form of electronically submitted reviews of hotels and electronic items (Salvetti 2014). While this in itself certainly is interesting material, it is hardly useful in the analysis of uh/um hesitation and pauses in spoken dialogue. Most studies concerning deception therefore generate their own source material by eliciting lies in a lab-setting. Newman et al. (2003) for example asked participants to explain their views on abortion to an interviewer once in a truthful manner and once in a deceptive manner (Newman et al 2003: 667). However, as this paper will later explain in more detail lab-generated lies tend to be unreliable representation of actual lies. The general issue with this is that the speaker is under little pressure to succeed with his act of deception and therefore puts considerably less effort into the production of the lie (for an example see Ekman 1985/1992: 54), making it hardly representative of the actuality of deception. It was therefore deemed necessary to focus on naturally occurring lies in which there is proper motivation for the speaker to conceal their agenda. To find these "naturally occurring" lies it made sense to look for highly publicized cases in which a considerable amount of truth-telling and lie-telling footage is accessible to the public in the form of interviews or deposition tapes. In addition to that, in order to have data than can be adequately compared between the truthful speech scenario and the deceptive speech scenario it was also deemed necessary to have both types of statements to be produced by the same person. This is due to the fact that people tend to have vastly differing speech patterns, employing hesitation and pauses at different frequencies and durations (Maclay and Osgood 1959: 33-36, Tottie 2011: 192, Wieling et al. 2016: 15). Therefore, this paper found it impractical to compare source material of different cases to one another and chose to focus on one individual, that being Holmes. Holmes' statements therefore are an ideal set of data as her lies are 1) naturally occurring and 2) highly motivated by financial and personal gain. Furthermore, since Holmes' became a very public

figure in the mid 2010s there is a considerable amount of footage that can be used for the analysis.

4. The Linguistics of Deception

4.1 Deception in Speech Acts

Before delving deeper, this paper would like to discuss the specific type of lie which will be investigated in the following analysis.
In its arguably simplest definition, a lie is: "something you say that you know is not true" (Cambridge Dictionary 2020a). Yet quantifying the status of lying in the context of traditional linguistics has proven to be rather difficult. After all, the act of lying is essentially a violation of Grice's (1975) cooperative principle by ignoring the conversational maxim of quality, where the speaker commits to speech acts that are deemed subjectively truthful (Fallis 2019: 187, Stokke 2019: 146). Traditionally, such violations have been identified by Grice (1975) as implicatures, i.e. utterances with implied meaning that despite violation of one or more maxims still serve conversational function (Grice 1975: 26, 30-31), examples of these being irony and hyperbole (Grice 1975: 34-35). However different from implicatures, a lie aims to conceal the transgression of the maxim from the listener and function as if no violation has occurred (Ortony and Gupta 2019: 149).

Suppose the following situation: A man meets a friend in a random encounter. The man has had a particularly difficult day but does not wish for his friend to be aware of his situation. They have the following exchange:

> Friend: Hey! How are you doing today?
> Man: I'm fine!

In a truthful speech scenario, the utterance "I'm fine!" can easily be defined by Austin's (1962) three part speech act system. The speaker intends to convey his well-being in the form of an assertion (illocutionary act), the assertion "I'm fine!" is produced (locutionary act) and the listener perceives the messages as intended, meaning he is informed of the speaker's well-being (perlocutionary act). However, in this untruthful speech scenario, though the

locutionary act and the perlocutionary act are consistent, the intention behind the illocutionary act changes. The intention is no longer to solely convey an assertion of one's own well-being, but to trick the listener into believing the assertion to be true despite the speaker's knowledge of it being untrue. In doing so the assertion "I'm fine!" no longer fulfills the demand of the "sincerity condition", which according to Searle (1969) is a needed preposition for a speech act such as an assertion to adequately fulfill its intention (Searle 1969: 64, 66). Yet, even with this preposition unfulfilled, the speech act is still successful in conveying the message. It has been therefore recognized that these types of "white lies" serve an important function within politeness theory. Bloomquist (2010) has for example stated that "in a social situation, politeness conditions override the truthfulness constraints, especially since social lies are meant to be helpful and flattering, not harmful." (Bloomquist 2010: 1602). One can therefore argue that there is little value in trying to uncover such "harmless" lies. What this paper is concerned with therefore, are not "white lies" but those lies that can be classified as deception. Defined as "the act of hiding the truth, especially to get an advantage" (Cambridge Dictionary 2020b) Bryant (2008) posits that "real lies" such as deception differ in the sense that the lie does not serve an underlying social function but instead serves the speaker and his agenda, even if it comes at the harm of the person being deceived (Bryant 2008: 32-33, Carson 2019: 478). This is defined by the motivation to "attain a gain (referred to as promotion focus) or to avoid a loss (prevention focus)" (Ortony and Gupta 2019: 150). So, while both deception and white lies function on the basis of the speaker hoping to prevent the listener from realizing that one if not several of Grice's (1975) cooperative principles have been broken (Ortony and Gupta 2013: 149) deception promises gain beyond face-saving strategies.

Taking one of the most famous lies as an example, Bill Clinton's response to the Lewinsky allegations on January 26, 1998 illustrate lying and deception taking place.

> But I wanna say one thing to the American people. I want you to listen to me. I want to say this again: I did not have sexual relations with that woman. Miss Lewinsky. I never told anybody to lie. Not a single time. Never. These allegations are false. And I need to go back to work for the American people. Thank you.
>
> (Bill Clinton 1998 in: APArchive 2020)

The obvious outright deceptive lies here are that 1) Clinton did not have an affair with Lewinsky and 2) Clinton did not engage in a cover-up of the affair, when by now the public

knows both of these statements to be false. Unlike a white lie, these lies, if uncovered, directly harm the people the deception is targeting, in this case by destroying voter trust and potentially compromising state security. Similarly, Holmes' act of deception promised her financial gain, social status and networking opportunities at the cost of Theranos' patient's safety and well-being.

This is where the importance of uncovering a lie ahead of time comes into play. So how does speech serve as indicator for lying?

4.2 uh/um Hesitation and unfilled Pauses as a Sign of Deceit?

The prevailing theory is that the act of deception requires an increased amount of cognitive and emotional energy to be expended by the speaker, as the speaker not only has to monitor his narrative and choice of words but also control any emotion like fear, guilt or delight the deception may invoke in them (Vrij et al. 2008: 38-40)[3]. Ekman (1985/1992) here defines the term *detection apprehension* as the general pressure a liar feels when confronted with the stakes of their deception (Ekman 1985/1992: 211). With higher detection apprehension leakage of cues tend to become more probable (Ekman 1985/1992: 92, 211). The argument being, an individual that suffers under the implications of stress, and with that cognitive load, is more likely to exhibit cues that lead to detection (Vrij et al. 2008: 396-397).

Independently from this, linguists have found evidence suggesting that cognitive load is largely responsible for a number of speech disruption, including both filled and unfilled hesitation pauses. Early on in the research of hesitation phenomena, Lounsbury (1954) had posited that "hesitation pauses correspond to the points of highest statistical uncertainty in the sequencing of units of any given order" (In: Osgood et al. 1954: 99). Building on this hypothesis Maclay and Osgood had found that hesitation pauses are indeed more likely to occur at instances of lexical decision making, as they had found hesitation to occur more frequently preceding lexical words than function words (Maclay and Osgood 1959: 32-33). Furthermore, they had found filled hesitation pauses to occur predominantly at the phrase boundaries as opposed to other locations within sentence structure, i.e. word boundaries (Maclay and Osgood 1959: 34). Similar to that, Boomer (1965) had found hesitation to occur at the boundaries of phonemic clauses and shortly after their initiation after the first word uttered (Boomer 1965: 155-156). These results were later confirmed by further evidence by

[3] For more detail see: Vrij et al. 2008: 43-46

Clark and Fox Tree (Clark and Fox Tree 2002: 93-95). This in turn leads to the idea that hesitation is not only due to cognitive processes involved in the lexical decision-making stage but is also due to efforts made concerning structural planning. Today, it has become apparent that hesitation markers and pauses are caused by a number of different factors and are therefore generally representative of "some kind of problem" within speech-flow (Tenbrink 2020: 136). Tenbrink (2020) names here not only before mentioned cognitive load but also issues with speaker conviction, general uncertainty and communicative concerns as further factors (Tenbrink 2020: 136). So, whilst not all hesitation occurs due to complications within speech production, most conditions require the speaker to be under increased stress one way or another to disrupt the regular speech flow. These findings should therefore suggest that a deceptive speaker, as he is suffering the consequences of increased cognitive load and general stress, is prone to utilize more hesitation pauses within his speech.

However, despite this, current research on the subject is highly contradictory.
First one needs to look at the issue of filled pauses, and with that uh/um hesitation markers, within deceptive speech. Villar and Arciuli (2012) analyzed the speech pattern of a convicted murderer speaking in a formal setting (TV interview) and informal setting (telephone tapes). Isolating the truthful and deceptive statements made within these they found that the individual was less likely to produce "um" hesitation markers when lying during both instances (Villar and Arciuli 2012: 91). This had been done similarly before by Vrij and Mann (2001) who had come to the same conclusion (Vrij and Mann 2001: 193). However, in another study, Vrij et al. (2000) had analyzed the speech pattern of medical students (nurses) who were prompted to lie spontaneously about the contents of a video they had just watched. In these instances, Vrij had et al. (2000) found that the individuals included more hesitations within their language than when they were speaking truthfully (Vrij et al 2000: 250-251). Again, a similar experiment conducted by Buller et al (1994) using interview samples from Bavelas et al. (1990) had found that truthful speakers indeed used less filled pauses than speakers uttering statements they determined as equivocate (Buller et al 1994: 414).
These paradoxical results are representative of the two biggest meta-analysis done on the subject of deception analysis so far. DePaulo et al. (2003) compiled several studies accumulating to 1338 estimates of 158 cues to deception suggesting linguistic, paralinguistic and physiological shifts within deceptive speech situations and then cross-referenced these in order to establish any significant overlap in findings. In doing so they made the following discoveries for what they defined as 'filled pauses' "ah", "er", "uh", "um" and "hmmm"

(DePaulo et al. 2003: 114). Initially, the use of filled pauses decreased in studies where the deceptive individual was motivated to succeed at lying, with an even stronger decrease when that motivation was of instrumental nature (e.g. financial) (DePaulo et al. 2003: 97). However, when that motivation was absent the deceptive speaker actually increased their use of filled pauses beyond the regular frequency (DePaulo et al. 2003: 97). It is important to note here that though these slight changes were noted DePaulo et al. (2003) did not deem these to be statistically significant enough to fully rely on them as indicators for deception (DePaulo et al. 2003: 97). The findings within Vrij et al. (2008) meta-analysis were similar in the sense that the selected studies showed drastically different results in hesitation marker production individually (Vrij et al 2008: 55, 90-94), but when cross-referenced showed that there could be no correlation made between hesitation markers and deception (Vrij et al 2008: 94, 124). Vrij (2008) therefore deemed hesitation markers to not be reliable enough to be indicators for truth- versus lie-telling (Vrij et al. 2008: 55). They had suggested that combining several cues instead leads to more reliable results (Vrij et al. 2008: 66).

Pauses, on the other hand, do show significant change from their regular pattern within deceptive statements. This paper would like to make clear here that when referencing pauses the so-called inter-lexical pauses, i.e. those that occur between words within sentence structure as opposed to intra-segmental pauses which occur within words at the phoneme level, (Zellner 1994: 42-43) are being talked about. This does not include any pauses occurring within the latency period of speech production or pauses that are due to breathing. This paper is therefore only concerned with pauses exceeding >= 200ms in length, the reasoning as to why being further discussed in Data and Methodology.

Vrij et al. (2008) found that liars include longer pauses within their utterances than truthful speakers, however the frequency of said pauses does not vary noticeably between speech scenarios (Vrij 2008: 55). Vrij et al. (2008) posits that this is indicative of lying being "somewhat more cognitively demanding than truth telling" (Vrij et al. 2008: 57). DePaulo et al. (2003) had found that speakers whose lies were of an identity-related nature were more likely to employ pauses than a truthful speaker would (DePaulo et al. 2003: 97). However, these findings again were not deemed statistically relevant, though they came close (DePaulo et al. 2003: 97). DePaulo et al. (2003) furthermore found that speaker's who could not plan their lies ahead of time were inclined to include more pauses (DePaulo et al. 2003: 100, 103).

4.3 uh/um Hesitation: Symptom or Interjection?

Lastly, before this paper delves into the research it is important to explain why the function and linguistic status of uh/um hesitation marker is still being debated and how research similar to this could potentially help in uncovering more evidence for the importance of hesitation markers within language.

In the early days of linguistic research, hesitations had been categorized as part of the larger phenomena of speech disfluencies. Not yet being concerned with their potential significance, Chomsky (1965) had in turn quantified such occurrences as "errors (random or characteristic) in applying knowledge of language in actual performance" (Chomsky 1965: 3). The argument implied speech errors not being a part of language itself but a problem in cognitive processes that hinder the production of language. It was therefore not relevant to early linguistic theory. With an advance in the linguistic field later research however, indicated a shift in the perception of disfluencies and started to attribute greater meaning to these seemingly negligible linguistic items. According to Levelt (1983) there is reason to believe some speech disturbances, including hesitations, to be part of language repair functions for the speaker. Levelt (1983) defines here the usage of "uh" hesitation markers as a way of signaling that trouble within the speech-flow has been detected by the speaker and that they are about to correct themselves accordingly (Levelt 1983: 41, 74-75). However, Levelt (1983) describes "uh" as a mere symptom of the "actuality or recency of trouble" and does not recognize hesitation markers to be part of language proper (Levelt 1983: 73). There were however others disagreeing with this hypothesis and claimed hesitation markers to have lexical properties: Earliest suggestions came from James (1972) who suggested hesitation markers to be interjections of the same kind as *oh, well, ah, etc*. yet James (1972) did not build a further hypothesis upon that claim. Building on these foundations Clark and Fox Tree (2002) argue hesitation markers to indeed be interjections with consistent phonology, prosody, syntax, semantics and pragmatics (Clark and Fox Tree 2002: 103-104). Clark and Fox Tree (2002) claim the speaker to have active control over the production of uh/um, which would be contradictory to the idea of hesitation markers being mere symptoms of cognitive load (Clark and Fox Tree 2002: 104-105). They argued here that uh/um hesitation markers "have conventional forms and meanings, conform to the notion of word syntactically and prosodically and contrast with another signal of delay, the process of prolongation" (Clark and Fox Tree 2002: 105). They had furthermore posited that uh and um differentiate slightly within their function and argues uh is uttered to signify small delays in speech production

whereas um indicates longer delays in speech production (Clark and Fox Tree 2002: 104). Villar and Arciuli (2012) who had found that uh/um hesitation occurs less frequently in the deceptive speech acts of a convicted murderer had argued for Clark and Fox Tree's arguments. They had posited that since the speaker in question employed less hesitation in the deceptive speech acts he may have been purposefully altering his speech output in order to deceive the listener (Villar and Arciuli 2012: 92), lending credence to the idea that uh/um hesitation production can be actively controlled. However, they also noted that further research on the subject needs to be conducted (Villar and Arciuli 2012: 92).

4.4 uh/um Hesitation: Location

Though the lexical status of hesitation markers remains a controversial subject, the location of hesitation markers in sentence structure has been more thoroughly explored and offers insight into how hesitation production relates to speech planning.

As stated before Maclay and Osgood (1959), Boomer (1965) and Clark and Fox Tree (2002) had all found hesitation to predominantly occur at the boundaries of units of relating information. While in Boomer's (1965) and Clark and Fox Tree's (2002) case these were phonemic clauses (Boomer 1965: 155-156) and intonation units (Clark and Fox Tree 2002: 93-95) respectively, Maclay and Osgood (1959) had found hesitations to occur predominantly at phrase boundaries whereas pauses tend to occur within syntactic units (Maclay and Osgood 1959: 39) giving evidence to the claim that hesitation does not occur arbitrarily but corresponds to locations of increased cognitive activity due to speech planning efforts. They had furthermore found hesitation to be predominantly located preceding lexical words as opposed to function words (Maclay and Osgood 1959: 39). Kjellmer (2003) too had found more detailed evidence supporting the claim that hesitation occurs mostly at the boundaries of phrases. Kjellmer's (2003) research shows that though noun phrases and adjective phrase are usually internally uninterrupted by hesitation within the phrase (Kjellmer 2003: 174-175) verb phrases and preposition phrases tend to incorporate more hesitation within themselves at the word boundary (Kjellmer 2003: 176-178). Furthermore, the same research had found hesitation tends to appear at the conjunction between coordinating and subordinating clauses (Kjellmer 2003: 180) with most collocations at the L1 location being the conjunction "and" and collocations at the R1 position being the personal pronoun "I" (Kjellmer 2003: 173). Earlier research by Quinting (1971) supports claims of this distribution as he had also found hesitation to favor preceding subject personal pronouns resulting in a majority of hesitation to

be situated pre-noun or pre-noun phrase (Quinting 1971: 59). Though this goes against Maclay and Osgood's (1959) claims of having found hesitation predominantly succeeding lexical words as opposed to function words, which would include personal pronouns, there seems to be at least a general consensus as to where hesitation occurs in the overarching sentence structure, i.e. at the boundary of units encoding relating information.

5. Predictions/Hypothesis

Confronted with this avalanche of contradictory results it is difficult to establish a solid hypothesis for the results of the analysis, yet the specific context of the analysis subject allows for some predictions to be made.

5.1 Hypothesis I

I. uh/um hesitation markers will occur slightly less frequently within the deceptive statements than within the truthful statements.

This is based on the idea that individuals who had time to prepare their lies ahead of time are less likely to be hesitant during conversation, as they do not have to devout energy to constructing a narrative on the spot (Miller et al. 1983: 114-115). Being invested into the same fraudulent scheme for several years (2003 - 2018) it is appropriate to assume Holmes to not only be versed in telling lies but also having constructed a certain set of deceptive answers to common question that may be asked about Theranos. In addition to that, there is a chance Holmes was informed about the questions of the interviewer before coming onto the show which would give her further opportunities to calculate her lies in advance. Furthermore, as DePaulo's et al. (2003) research indicates, liars tend to use less hesitation markers when there is ample motivation for them to avoid apprehension, especially of the financial kind (DePaulo et al 2003: 97). As failure of her deception would cause Holmes' not only to lose her reputation and business but also her freedom, Holmes' incentive to succeed is immense. However, as DePaulo et al. (2003) had found these findings to be statistically insignificant the assumption is that this research will also show only marginal differences. As Vrij et al. (2008) had also commented on the fact that a prepared liar could show less signs of cognitive load (Vrij et al. 2008: 151) this could furthermore signal a slim decrease in hesitation marker usage.

5.2 Hypothesis II

II. Pauses will be of longer duration during Holmes' deceptive utterances compared to her truthful ones. Pauses will however not be more frequent within deceptive utterances than within truthful ones.

This assumption can be made based on the findings made by Vrij et al (2008) who had come to that exact conclusion (Vrij et al. 2008: 55, 84, 90-93). It is important to reiterate here that DePaulo et al (2003) have not found pauses to be relevant for deception detection, though they did find pauses to be more frequent when the motivation for deception was identity-related (DePaulo et al 2003: 97). Yet DePaulo et al (2003) only focused on the frequency of pauses and neglected to account for the duration of pauses.

5.3 Hypothesis III

III. uh/um hesitation markers will not occur arbitrarily throughout the utterances, but there are points in sentence structure where they are statistically more likely to occur. These points are presumably the phrase boundaries.

Clark and Fox Tree (2002) and Maclay and Osgood (1959) had found hesitation markers not to occur arbitrarily and to be most commonly placed at intonation unit boundaries (Clark and Fox Tree 2002: 93-95) and phrase boundaries respectively (Maclay and Osgood 1959: 39). Kjellmer (2003) had supported this partially by finding hesitation both preceding as well interrupting phrases depending on which kind of phrase is being looked at (Kjellmer 2003: 175-179). Though these units do not always overlap their research does seem to suggest that hesitation seems to occur before groups of words representative of one unit of "thought" as Kjellmer (2003) described it (Kjellmer 2003: 174). Since Kjellmer's (2003) approach is the most extensive and conclusive one to date, this paper will follow his structure of analysis. In doing so this paper will analyze for hesitation location on three levels: 1) Clause level, 2) Phrase level and 3) Collocation level, in an attempt to either find evidence or counter-evidence in regards to previous claims by the before mentioned authors. The current expectations are that Holmes' hesitation marker usage will adhere to Kjellmer's (2003) findings, meaning that both uh/um will be found predominantly at conjunctions whilst

preceding mostly the noun phrases of the following coordinating/subordinating clause. The collocations are therefore expected to be conjunctions such as “and” in the L1 position and pronouns such as “I” at the R1 location.

6. Data and Methodology

6.1 Participant

As described in 2. Context the participant of this study is Elizabeth Holmes. Holmes is a North American female and ranged between the ages of 25 to 31 during the interviews relevant to this analysis. Holmes was born on February 3rd 1984 in Washington D.C. but lived in Silicon Valley from 2002 onwards. Her highest level of academic achievement is a high school degree however Holmes did attend Stanford University in 2002-2003. Holmes’ first language is English. She has no known speech impediments, though it has been alleged she artificially lowers her voice when speaking to sound more serious. This accusation has not been confirmed.

6.2 Data

Source for the following analysis will be six different interviews given by Elizabeth Holmes to various media outlets and institutions, those being: the Stanford Technology Ventures Program (2009), Bergman Klein Centre (2015), Vanity Fair (2015), CNBC (2015), Fortune Magazine (2015) and TechCrunch Disrupt (2014). Henceforth these interviews will be called Interview I-VI based on the order given here. The interview audio footage is taken from the media outlet’s respective YouTube account.

6.3 Procedure and Programs

The interviews were transcribed according to DuBois et al. (1993). Here details that are not relevant to the analysis such as shift in pitch or lengthening are not part of the transcription. Prior to the analysis each transcript was again compared to the audio to ensure their quality. Each transcript was then uploaded as a txt file into AntConc (Anthony 2020), which was used to establish a concordance profile and determine collocation in the L1 and R1 location. From

this Bloomberg's (1933) and Wells (1947) approach to IC-analysis was used to analyze for the hesitation markers position within clause and phrase structure.

When it comes to the analysis of pauses, the issue presents itself that most research either uses programs that are not publically available or ones that are coded by the research team in question. This paper therefore found it necessary to find a publically available sound editing program which could perform the same procedures. This paper chose Audacity (Audacity Team 2020) for the analysis. The audio track is then uploaded into the program. If needed the file is treated with a background noise equalizer to eliminate potential microphone static interference or audience chatter. The interview at the Berkman Klein Centre (2015) was at some points disrupted due to a microphone malfunctioning. No statement given during such a malfunction will be used except for Interview II Statement b where the malfunction was minimal and easy enough to repair in Audacity. For the pause detection Audacity's automatic pause finder is then set to mark the file for pauses >=200ms at <=30dB. The threshold of 200ms is chosen as a conservative approach to the issue of ongoing disagreement between researchers when it comes to the distinction between articulatory pauses and hesitation pauses. The main problem arises when trying to establish a cut-off point for pause detection. For the longest part the golden standard of pause detection was set by Goldman-Eisler (1968) at silence exceeding 250ms in duration. Goldman-Eisler posited speech pauses <250ms to be of solely 'articulatory' nature with no connection to hesitation phenomena (Goldman-Eisler 1968: 12). While a significant amount of research uses cutoff points around this margin (150-250ms), some have criticized this threshold as being too lax. Hieke et al. (1983) considered short pauses with a duration frame of 130-250ms to indeed be relevant outside of articulation and suggested a minimum of 100ms to be usable in experiments (Hieke et al 1983: 212). For lack of unanimity within previous research this paper will use a more conservative threshold of 200ms, as not accidentally include physiologically caused pauses in the analysis. The second threshold of 30dB is chosen based on preliminary testing with each audio file, which showed Audacity to measure most accurately at that frequency. To further diminish the potential for inaccuracy Audacity's automatic sound finder is also used to highlight areas of sound over areas of silence (>=200ms, 30dB). Both silence and sound finder are overlapped to cross-examine their accuracy.

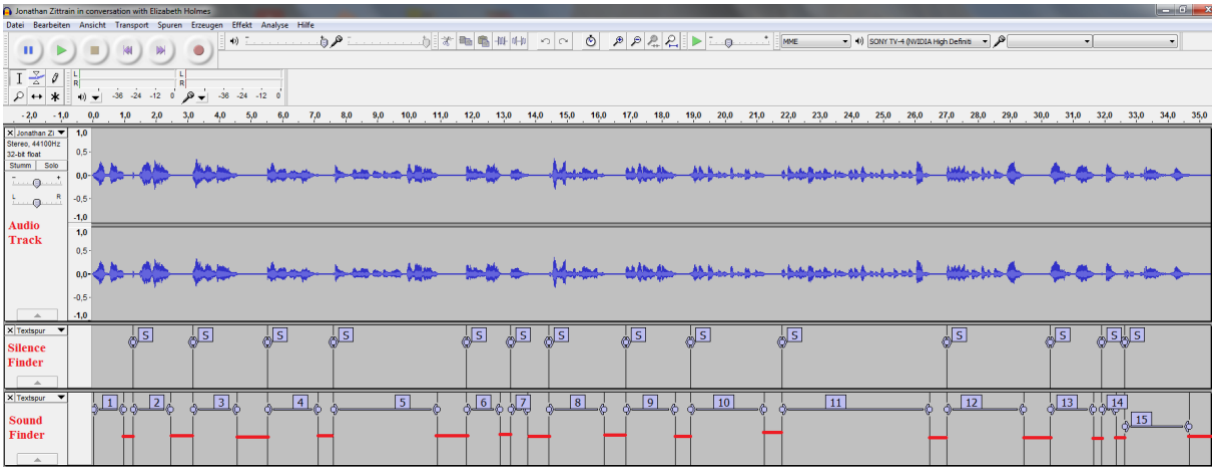


Figure 1.0: First 35.5 seconds of Holmes' statements in Interview II (Berman Klein Center 2015) in Audacity
Red lines in the Sound Finder bar illustrate where the program picked up on pauses >=200ms, 30dBz.

6.4 Defining Truth and Lie

The given interviews can be divided into two categories: truthful and deceptive, the definition of which follows those employed by Vrij and Mann (2001), Vrij and Mann (2002), Davis et al (2005) and Villar and Arciuli (2012). In the *truthful* interviews, Holmes' recounts her early education and time at Stanford as well as events in her early adolescence which inspired her to create Theranos. These isolated statements can be considered truthful as they have been verified by other parties involved (e.g. Holmes' former Stanford professors) and have not been further contested.

Example:
"I started a company when I was in high school that was distributing C++ compilers to universities in China."

(Interview II, Statement b)

This statement is true. Holmes' did in fact have said business in high school.

Any statements that are of a highly subjective nature, e.g. personal opinions or description of emotions are excluded as their veracity cannot be sufficiently established. For example, during one interview Holmes states: "one of the core elements […] of why people don't get access to tests is fear of needles". First of all, this statement can be considered factually wrong as many other factors such as costs and location also contribute to the lack of accessibility of medical treatment. Secondly, there is simply no way of confirming if Holmes' truly believes

in her statement or if she lied about it to make her product seem more appealing[4]. One minor exception to this rule of exclusion are Interview II, statements c & d in which Holmes explains personal reasons for the founding of Theranos. Many of her former employees and board members have stated that they believe this statement to be true and have cited her strong personal involvement as an issue as to why Holmes' would not want to cease operation of Theranos when issues with the product were raised to her.

Within the *deceptive* interviews Holmes' gives statements that have been proven to be a lie, either by Theranos Whistleblowers, the investigation by the Wall Street Journal or the SEC. Some minor parts of the statements contained within these interviews cannot be classified as outright lies but are more of an equivocate nature. These are still included in the analysis. However, just as within the truthful statements comments of a subjective kind are excluded.

Lie:
"[…] there is no reason we can't do peer-review and there is no reason we can't publish other stats and we're gonna do that."

(Interview V, Statement c)

This statement is a lie. Holmes was aware that Theranos data is falsified and unreliable and would not withstand being peer-reviewed by researchers in the medical field. Theranos never published any work to be peer-reviewed following this statement either.

[4] The Edison machine uses a very small needle and can therefore be considered less painful than traditional blood-testing equipment.

7. Results

7.1 uh/um Hesitation Frequency and Distribution Across Statement Length

The following are the results from the analysis of Holmes' statements given within the two truthful interview segments and the four deceptive interview segments.
First the research takes a look at the frequency of uh/um hesitation across all truthful and deceptive speech scenarios.

Truthful Speech Scenarios

	Interview I	Interview II
Words	194	515
Uh	4	5
Um	5	23
Time	1 min 28 sec	3 min 45 sec

Table 1.0: Word and Hesitation Count for Interview I and II

	Interview I	Interview II
Speech Rate [words/min]	132.0	137.3
Hesitation Makers per 100 words	4.6	5.4
Hesitation Markers per Minute	6.1	7.5

Table 1.1: Individual Averages for Interview I and II

Within the two truthful speech scenarios Holmes' utters a total of 709 words over the span of 5 minutes and 13 seconds combined. Within that timeframe she uses the hesitation marker "uh" a total of 9 times, whereas "um" is used a total of 28 times, resulting in a total of 37 hesitation markers being employed overall. The overall speech rate is ~136 words per minute. The amount of hesitation markers results in an average of ~5.2 markers uttered per 100 words and ~6.9 markers used per minute.

Deceptive Speech Scenarios

	Interview III	Interview IV	Interview V	Interview VI
Words	230	280	230	179
Uh	4	1	1	2
Um	2	3	2	1
Time	1 min 30 sec	1 min 34 sec	1 min 6 sec	1 min 11 sec

Table 1.2: Word and Hesitation Count for Interview III-VI

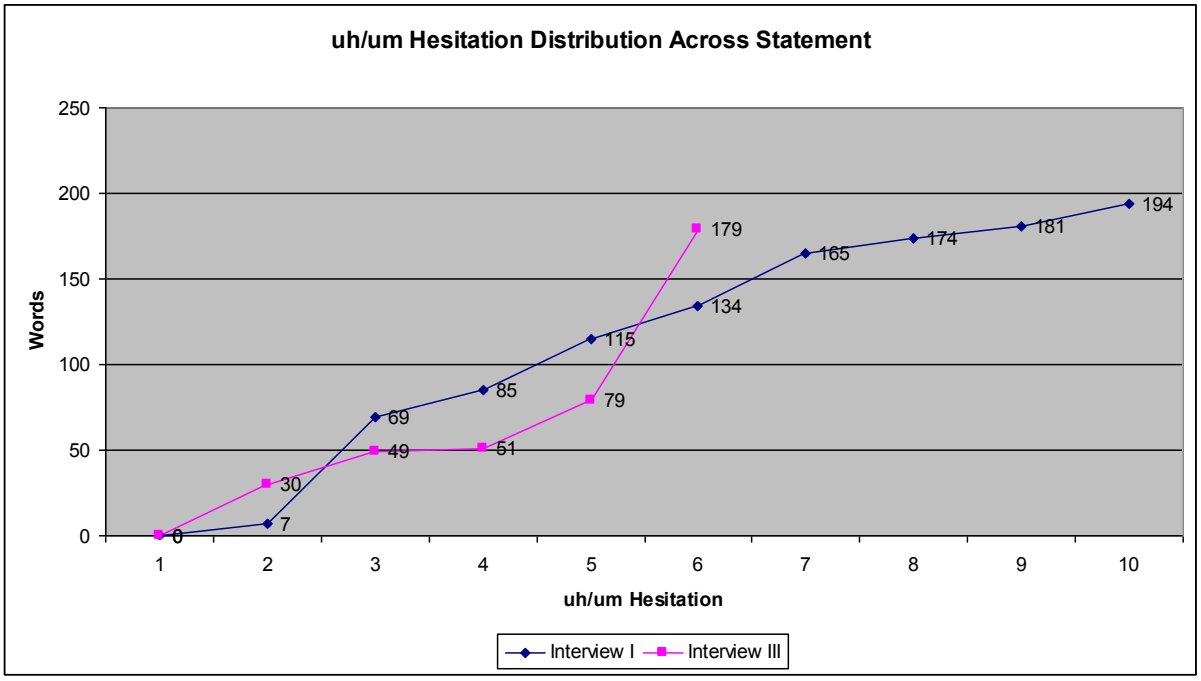
	Interview III	Interview IV	Interview V	Interview VI
Speech Rate [words/min]	153	178.7	209.1	151.3
Hesitation Makers per 100 words	2.6	1.4	1.3	1.7
Hesitation Markers per minute	4	2.6	2.7	2.5

Table 1.3: Individual Averages for Interview III-VI

Within the four deceptive speech scenarios Holmes' utters a total of 919 words over the span of 5 minutes and 21 seconds combined. Within that timeframe she uses both the hesitation marker "uh" as well as "um" a total of 8 times respectively, resulting in 16 markers overall. The average speech rate is 171.8 words per minute. The amount of hesitation markers results in an average of ~1.7 markers uttered per 100 words and ~3.0 markers used per minute.

This shows that uh/um hesitation occurs at only 43.2% the rate within Holmes' deceptive speech that it would in her truthful speech. While uh makes up 24.3% and um makes up 75.7% of all hesitation markers in truthful speech, both uh and um are equally distributed in deceptive speech (50% each). The drop in hesitation frequency in deceptive speech is rather surprising, reasons as to why being explored later (see 8. Discussion). In total Holmes' only produces 53 hesitation markers over the course of 10 minutes and 34 seconds of speech. What is important to note here too is that Holmes' speech rate is slightly elevated within the deceptive speech scenario. Here, Holmes produces 35.8 words per minute more than within her truthful speech pattern, meaning within truthful speech she only speak at about 70.2% the rate at which she does in deceptive speech.

Though the uh/um hesitation marker frequency in both truthful statements and the four deceptive statements is relatively similar within their respective category the distribution of the hesitation marker across the length of the statement seems rather random. When looking at each individual statement one can see that the neither uh nor um hesitation clusters somewhere within the statements, as in there is not majority of hesitation situated at the beginning, middle or end of the utterance. This lack of correlation exists both within the truthful statements and deceptive statements respectively as well as in comparison to one another. Taking a look at the distribution of uh/um hesitation markers in the truthful statement given in Interview I and in the deceptive statement given in Interview III as an example, one can see that though both utterances show a rather even distribution of hesitation across statement length, they still do not occur at the same points throughout (see Graph 1.0).



Graph 1.0: uh/um Hesitation Distribution Across Statement Length

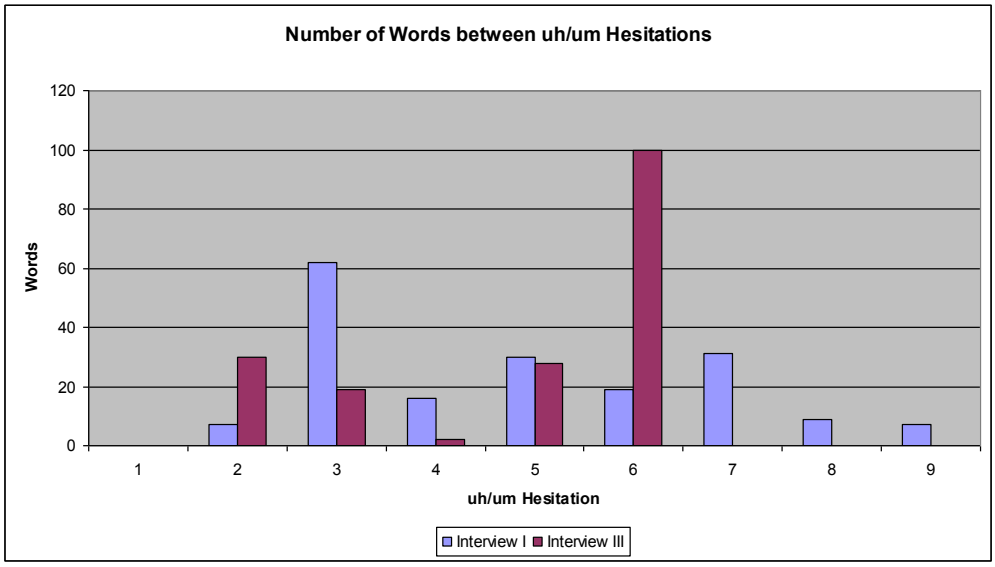
To further illustrate this using the same example of Interview I and Interview III, one can count the amount of words that pass between each hesitation occurring respectively in each interview (see Graph 1.1). Again, no correlation can be established. There is no average amount of words that have to occur between each hesitation before another can be uttered. Hesitation can arise as frequently as occurring multiple times within one clause with some hesitation even immediately following each other as in the second clause of the following

Example:

"I started **uh** a company when I was in **uh uh** high school **um** […]"

(Interview II, Statement b)

However, two hesitation markers can also be separated by long segments of uninterrupted speech. In Holmes' example the longest she has spoken without employing a hesitation marker lies at 107 words (Interview VI, Statement a).



Graph 1.1: Number of Words between uh/um Hesitations

This seems to suggest that no one part concerning the length of any given utterance is more likely or less likely to be interrupted by speech production errors. As with each new sentence across statement length new lexical decision making and structural planning has to be conducted it is unsurprising that that the employment of hesitation markers reflects this.

7.2 Pause Duration, Frequency and Distribution Across Statement Length

Truthful Speech Scenarios

	Interview I	Interview II
Number of Pauses	28	101
Average Pause Length	604ms	543ms
Total Pause Duration	16.920ms	54.860ms

Table 2.0: Pause Count and Length for Interview I and II

	Interview I	Interview II
Pauses/100 Words	14.4	19.6
Pauses/min	19.0	27.0

Table 2.1: Individual Averages of Pauses for Interview I and II

In the truthful speech scenarios a total of 129 pauses are implemented over the course of the 5 min and 13 seconds of speech duration. Of that time a total 1 min 12 sec are taken up by these pauses >=200ms. The duration of pure speech is therefore reduced to 4 min 1 sec. The

overall average pause length sits at 556ms. The combined average amount of pause per minute is 24.7.

The longest pause in the truthful speech segment is measured at 1.520ms.

Deceptive Speech Scenarios

	Interview III	Interview IV	Interview V	Interview VI
Number of Pauses	22	20	10	23
Average Pause Length	460ms	410.5ms	403ms	412ms
Total Pause Duration	10.120ms	8.210ms	4.030ms	9.880ms

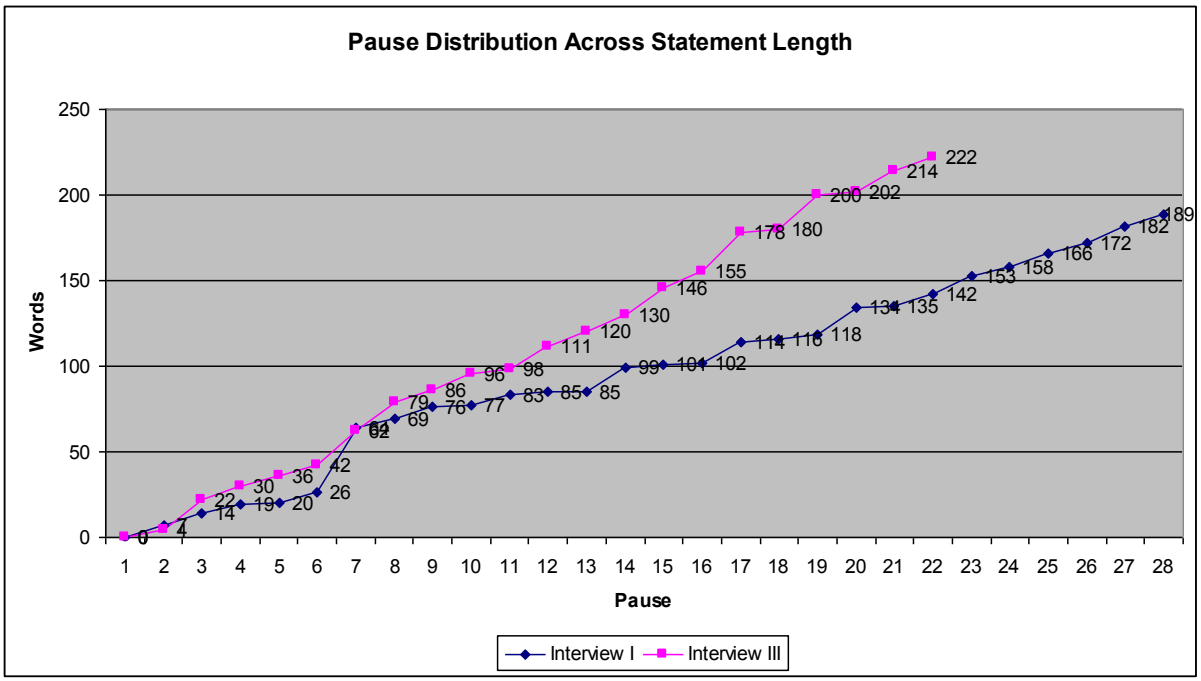
Table 2.2: Pause Count and Length for Interview III-VI

	Speech III	Speech IV	Speech V	Speech VI
Pauses/100 words	9.6	7.1	4.3	12.8
Pauses/min	14.7	12.8	9.1	19.4

Table 2.3: Individual Averages of Pauses for Interview III-VI

In the deceptive speech scenarios a total of 75 pauses are implemented over the course of the 5 min 26 sec of speech duration. Of this time a total of 0 min 32 sec are defined by pauses >=200ms leaving 4 min 54 sec of uninterrupted speech. The average pause length is 430ms. The combined amount of average pause per minute is 13.8. The longest pause recorded in a deceptive speech act is 970ms.

Using the same example presented in Graph 1.0 to represent the usage of pauses across statement length for the utterances within Interview I and Interview III one can see the following distribution:



Graph 2.0 Pause Distribution Across Statement Length

Again, just like uh/um hesitation markers pauses occur rather evenly distributed across statement length and do not cluster around the beginning, middle or end of any given utterance. As one can see however, pauses are more numerous and therefore happen with an increased frequency as opposed to uh/um hesitation. As a result of that fewer words are spoken between each pause than compared to uh/um. Pauses can occur as frequently as having only one word separating them. Though not included in the word count this "word" could even be an uh/um hesitation marker as in the following sentence:

"[…] originally the concept was well **P I P** take a a leave of absence **P** and then **P uh P** it became really clear that you know […]

(Interview I, Statement a)

However, just as with hesitation markers, there are also longer areas of speech production that remain uninterrupted by pauses, though these tend to be nowhere near as long as when it comes to uh/um. The longest period of speech that is uninterrupted by pauses in Interview I is about 38 words long. In Interview III the longest uninterrupted stretch of speech is only 23 words long.

7.3 Hesitation in Clauses

Before delving into the results of the distribution of uh/um hesitation across sentence structure this paper would like to acknowledge that due to the unexpected decrease in hesitation marker usage during Holmes' deceptive speech pattern the paper will have to work with a smaller number of hesitation markers, and therefore a smaller source pool for the coming analysis, than anticipated (53 in total). While the amount of hesitation markers is enough to establish a pattern of their occurrence within clause structure and phrase structure, this becomes more difficult when it comes to collocation distribution. This will be discussed in more detail in the corresponding section (see 7.5).

First, this paper wishes to take a look at how uh/um hesitation is distributed on a sentence/clause level. Kjellmer (2003) had only focused on the beginning of clauses (Kjellmer 2003: 180), meaning his research had only focused on those occurring at the beginning of a new sentence and those situated at the conjunction of two clauses. This paper however, would like to make sure that all hesitation is noted and will therefore use the following five categories to sort their appearance in:

1) Before a Clause: The immediate beginning of a new clause that has not been preceded by a conjunction.
2) Before a Conjunction: At the end of a clause and before the conjunction
3) In a Conjunction: At the end of a clause and between two conjunctions
4) After a Clause: At the end of a clause and after the conjunction
5) In a Clause: Neither at the beginning or the end of a clause but situated within

Category 3 warrants some explanation. In her utterances Holmes tends to repeat conjunctions as either a form of speech error or in an attempt to restart the following clause. In many of these cases she employs a hesitation marker between these two conjunctions which lead to utterances such as: "[…] and **uh** and […]". Holmes furthermore has a tendency to not use conjunctions such as "and" by themselves but extends them by using the adverb "then" or adverb/conjunction "so" leading to conjunction structures such as: "[…] and so **um** so then […]". In other instances in Holmes' speech "so" and "then" are used in ways in which they replace each other or the "and" at a coordinating point of clauses completely, leading them to fill the role of the conjunction, which in the case of "then" would be grammatically incorrect in written text, yet acceptable in spoken speech (i.e. saying just "then" instead of

"and then"). For these reasons this paper accepts combinations such as "and so then" as conjunctions consisting of several elements. The before mentioned example of "[…] and so **um** so then […]" will therefore be counted as hesitation occurring between conjunctions.

	BefClause	BefConj	InConj	AftConj	InClause
Uh	1	1	4	2	9
Um	9	9	4	9	5

Table 3.0: Hesitation in Clause Structure

In total 39 out of 53 hesitations (73.6%) occur at instances that signal the beginning of a new clause, with the remaining 14 being situated within clauses. Most of the hesitation (29, 54.7%) can be found in and around the conjunctive element of either coordinating or subordinating sentences. Here, 10 instances can be located before the conjunction, whereas another 11 are found immediately after the conjunction.

Example:

> "I think that was a really wonderful and special thing **um** and **um** I I er was very normal in the context of working to pursue something that that I loved."
>
> (Interview II, Statement a)

In this sentence, Holmes employs the um hesitation marker both before and after the conjunction "and", essentially signalling the end of the clause "I think that was a really wonderful and special thing" and the beginning coordinating clause "I was very normal in the context of working to pursue something that I loved." Holmes also employs a different hesitation marker "er" in the second clause. Again, as this paper is focusing on uh and um the "er" was not deemed relevant.

In eight of these cases the hesitation occurs between or "in" conjunctions. As before stated this is mostly the case when the conjunctive element is said twice due to speech repetition or a restart of the coordinating/subordinating sentence, but also happens when a conjunction is extended by a "so" or "then" or both. Arguments could be made that these should be counted towards the "after conjunction" section, yet since these occurrences are not isolated cases the paper wishes to highlight them within their own category.

Example:

> "I ended up living in the basement of a bunch of of Stanford grads and started working from the basement of their house and **uh** and then finally convinced someone to to give us **uh -** originally it was a bridge loan […]"
>
> (Interview I, Statement a)

Kjellmer (2003) had found within his research that hesitation is more likely to succeed conjunctions than it is to precede them (Kjellmer 2003: 180). He had argued that this is due to the speaker wishing to signal that he intends to hold the floor and is preparing a new sequence of utterances (Kjellmer 2003: 183). As before stated, since the repetition of a conjunction word is a speech error and would be deleted in a clean repeat of the sentence, any hesitation that sits between repeated conjunctions can be counted towards the "after conjunction" count. This combined count would result in 10 hesitation markers being positioned before a conjunction, whereas now a total of 19 hesitation markers can be found after the conjunction. For the few 14 instances in which as hesitation occurs not preceding a clause these are found within the sentence at various points throughout phrase structure.

Example:

"I was **uh** a young girl who loved science and math and engineering."

(Interview II, Statement a)

7.4 Hesitation in Phrases

Since conjunctions situated between clauses cannot be counted towards any type of phrase they are essentially removed from the following analysis. This means that when looking at a hesitation marker situated before, in between or after at a conjunction the phrase immediately following the conjunction will be the one of relevance. There are also three instances in which an um hesitation is categorized as preceding "other". This is the case when the hesitation marker was found preceding an interjection or a very short phrase that was abandoned due to immediate restart. Since these instances can hardly be classified within phrasal structure, they are being left out of the count.

	NP	VP	AdvP	PrepP	Other	WB
uh	12	4	1	0	0	0
um	24	6	2	0	3	1

Table 4.0: Hesitation in Phrase Structure

One can see here immediately that hesitation predominantly precedes noun phrases. This resonates with Quinting's (1971) before mentioned findings (see 4.4). Since most hesitation is situated at the beginning of a new clause this distribution seems logical. After all, as English is a SVO language, sentences are largely initiated by noun phrases. Interestingly enough, Verb Phrases, Adverb Phrases and Preposition Phrases are largely unrepresented within this. No Adjective Phrase could be found preceded by a hesitation marker.

Noun Phrase

Taking a closer look at the Noun Phrases there are essentially two types of common occurrences. As before mentioned since most hesitation occurs at the beginning of a new clause it is unsurprising to find these to be also preceding Noun Phrases.

Example:
"**um** I'd sort of convinced my way into working in one of the PhD programs […]"

(Interview II, Statement e)

In these instances the subject of the sentence, here signified by the subject personal pronoun "I", is the sole component of the initiating Noun Phrase of the sentence, which is then in turn followed by the Verb Phrase and further constituents. However, when occurring within the clause structure, the hesitation still tends to find itself at a Noun Phrase boundary, despite most sentences consisting of a variety of phrases.

Example: "I was **uh** a young girl who loved science and math and engineering."

(Interview II, Statement a)

Since the "who" in the given sentence serves to introduce the following relative clause the sentence can be divided up as "I was uh a young girl." And "I loved science and math and engineering.". Focusing on the portion containing the hesitation marker one can construct the following constituent phrase tree:

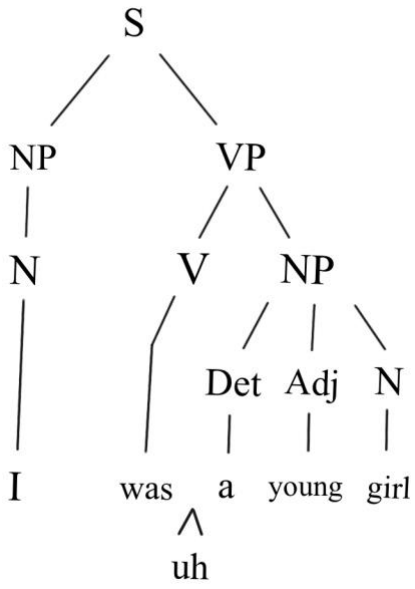


Figure 2.0: Syntax Tree 1

The uh hesitation marker is situated between the Verb Phrase of “was” and the Noun Phrase of “a young girl”. Examples such as these are interesting in the sense that there seems to be little indication as to why hesitation would not occur more frequently before Verb Phrases as well. As in why is it “I was **uh** a young girl” as opposed to “I **uh** was a young girl”? Arguments can be made that in examples such as these the production of the verb phrase is rather easy when it comes to both structural planning as well as lexical choice. The verb phrase here consists of only the main verb with no auxiliary verb supporting it. Furthermore, the linking verb “to be” does not offer many lexical alternatives as other verbs (main verbs/action verbs) may do. Therefore one can assume simple Verb Phrases as these are not cognitively challenging to construct and therefore warrant no production of hesitation markers as a reaction. Though also seemingly simple the following Noun Phrase of “a young girl” still offers more choice in lexical decision making and structural planning compared to the Noun Phrase “I” and the Verb Phrase “was” before it. The phrase could easily be replaced with a number of variants such as “a child” or “a little kid”. Therefore, if hesitation were to occur in the given clause of “I was a young girl” it stands to reason that the thought unit which is most difficult to produce and therefore may cause delay in speech production is the one initiated by a hesitation marker.

Verb Phrase

As before shown in table 4.0 Verb Phrases are not often preceded by hesitation. One rather straightforward example of these scenarios is the following:

"I had **um** family that **um** brought me up in such a way […]"

(Interview II, Statement a)

Here, the paper would like to highlight the on instance of hesitation in which the marker can be argued to be within a phrase instead of separating two phrases at their boundary. This occurs in the following sentence:

"[The company] that was **um** distributing C++ programming compilers to **um** universities in China"

(Interview II, Statement b)

This argument is based on contrasting views on the question of if auxiliaries and main verbs are to be separated in constituency structure or if such non-finite Verb Phrase can exist independently divided into their components. In the case of "was distributing" being part of one Verb Phrase which then divides into auxiliary and main verb this hesitation would indeed be interrupting the phrase at the word boundary:

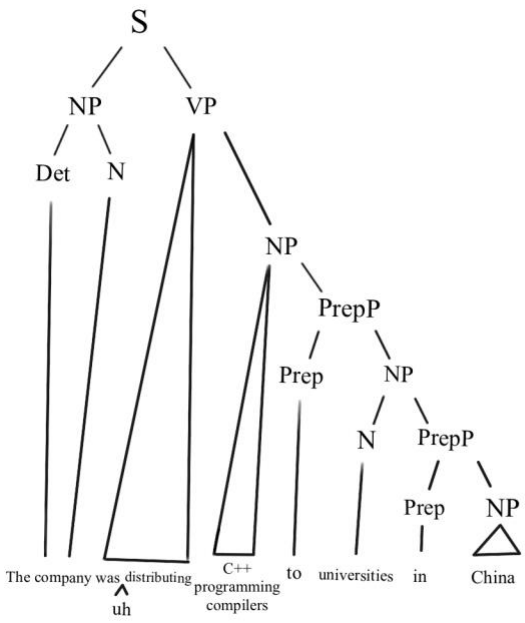


Figure 2.1: Syntax Tree 2

However, auxiliary and main verb to be divided into their respective Verb Phrases 1 and 2 one could argue the uh is not disturbing the speech flow on a word boundary level, but instead behaves like any other hesitation and sits at the phrase boundary.

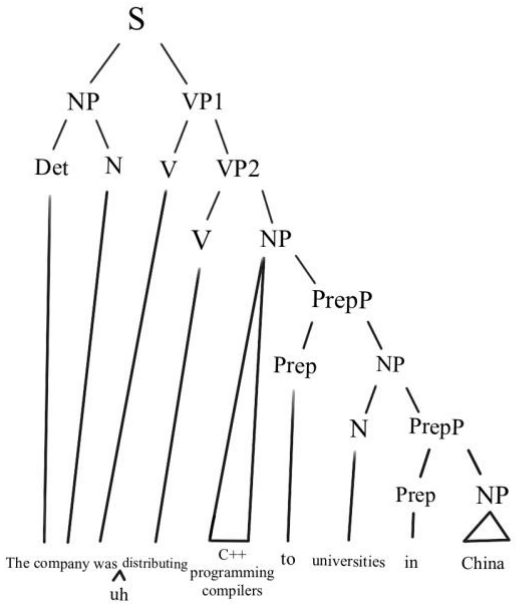


Figure 2.2: Syntax Tree 3

This highlights an issue with how differences in constituency approach by the researcher could alter data in regards to such hesitation placement. Kjellmer's (2003) approach does include seem to include auxiliary verbs within the same phrase as the main verb as he does so when it comes to the auxiliaries has/have/had in perfect tense and past perfect, though he had not specifically stated this to be also true for the past progressive as being used in this example (Kjellmer 2003: 176). Following Kjellmer's (2003) approach Figure 2.1 is the preferred version for this paper and therefore the uh hesitation in this instance does occur at the word boundary within a verb phrase.

Adverb Phrase

In the instances in which uh hesitation occurs preceding Adverb Phrases it is when the head adverb of the phrase is situated at the beginning of the clause as in:

"**uh** originally it was a bridge loan […]"

(Interview I, Statement a)

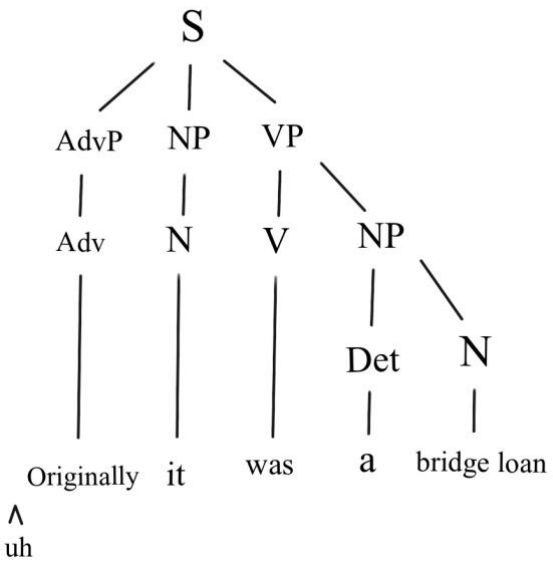


Figure 2.3: Syntax Tree 4

In these instances it is difficult to establish if the Adverb Phrase is actually the unit causing hesitation to occur or if the planning of the entire clause overall causes cognitive load. Since hesitation does not occur at Adverb Phrases within the clause this seems to suggest that uh/um is not employed due to errors in the production of the Adverb Phrase but more so due to issues with the construction of the entire upcoming clause, though this is difficult to verify.

Spoken Speech and Constituent Analysis

As one can see, analyzing spoken speech within the constraints of IC-analysis is rather difficult. Spoken speech does not behave as orderly as written text and therefore tends to break the rules of grammar in conversationally appropriate ways. Taking the following example one can illustrate this problem further:

"[…] **uh** and then finally convinced someone to to give us [a bridge loan]"

(Interview I, Statement a)

This clause is missing its subject/Noun Phrase which due to context is known to be "I". What is unknown however is where the "I" is to be placed as the sentence presents several options. The utterance could therefore be both "I finally convinced someone to give us a bridge loan" or "Finally, I convinced someone to give us a bridge loan". Whilst this does not have any implications for the delivery of the utterance it does for phrase analysis. Specifically, does the uh precede an invisible Noun Phrase "I", the Verb Phrase "finally convinced" or the

Adverb Phrase “finally”? Not wanting to alter the sentence further, even if it were to make it more grammatically correct, instances such as these were counted towards belong to the following Verb Phrase (adverb + verb). Unfortunately, neither Maclay and Osgood’s (1959) nor Kjellmer’s (2003) specified how they had treated incomplete clauses such as these. This paper therefore wishes to caution how such instances could influence the distribution.

7.5 Hesitation and Collocations

Uh Collocations	L1 Position	R1 Position
Noun	3	0
Verb	3	0
Adjective	0	2
Adverb	1	3
Pronoun	1	3
Preposition	3	1
Conjunction	5	4
Determiner	0	3
Uh Hesitation	1	1

Table 5.0: Uh Collocations

Both the L1 and R1 leaning collocations of uh hesitation markers are rather evenly distributed across word type, with a small increase at conjunctions being noticeable. This is due to the before mentioned statistical likeliness for hesitation to occur at the coordinating element of two clauses. However, since only 17 uh hesitation markers occur overall it is difficult to draw an accurate conclusion from such a small source pool without referencing the more numerous um hesitations.

Um Collocations	L1 Position	R1 Position
Noun	10	2
Verb	5	3
Adjective	1	0
Adverb	1	1
Pronoun	2	12
Preposition	1	1
Conjunction	13	15
Determiner	0	1
Interjection	1	1
[No Collocation]	2	0

Table 5.1: Um Collocations

Um hesitations behave similarly in the sense that most collocation at both the L1 and R1 position are conjunctions. Unlike uh hesitation this is however closely followed by pronouns in the R1 location (12, 33.3%) and nouns at the L1 location (10, 27.8%).

	Lexical Words	Function Words	Interjections
Uh	6	11	0
Um	6	28	2

Table 5.2: uh/um across Lexical Words and Function Words

Unlike the findings by Maclay and Osgood (1959) hesitation occurs predominantly preceding function words. This again is likely due to the dominance of hesitation at clause coordinating elements. Since most R1 leaning words are conjunctions and pronouns it is unsurprising that therefore majority of collocations are function words. Maclay and Osgood (1959) had hypothesized that hesitation occurs predominantly preceding lexical words, as these offer a greater degree of lexical decision making choices as opposed to function words (Maclay and Osgood 1959: 32), this being in line with the assumption that hesitation occurs at greater points of uncertainty (In: Osgood et al. 1954: 99). However, this idea is unsupported by the current data (for further comments see 8. Discussion).

7.6 Pauses and uh/um Hesitation Marker Interaction

	Uh	Um
Pause Before	5	13
Pause After	5	23

Table 6.0: Hesitation Collocation with Pauses

Out of a total of 17 uh hesitation markers about half (9/53%) are directly preceded or succeeded (or both) by silent pauses. Out of the 36 hesitation markers a total of 26 (72%) are preceded or succeeded (or both) by silent pauses. Whilst the distribution of pauses around the uh hesitation marker is equal when it comes to the position of before and after, the um hesitation marker is largely followed by a pause (64% of the time) as opposed to preceded by it (36% of the time). Clark and Fox Tree (2002) had argued that the hesitation marker um signals a longer delay in speech production than uh does (Clark and Fox Tree 2002: 104). Since this paper had found um to be indeed more likely to be accompanied by silent pauses these findings seem to support Clark and Fox Tree's (2002) claims. Knowing that hesitation markers are likely to be surrounded by pauses could also be further support for the claim levelled by Maclay and Osgood (1959) that hesitations tend to serve as a means to hold the floor in turn-taking strategies (Maclay and Osgood 1959: 41). In any conversation most points of silence present opportunities for the speaker's counterpart to take over the floor (Heldner and Edlund 2010: 556, 566). The utterance of uh and um therefore represents a means for the current active speaker to signal that they are not yet finished with their production of speech, yet struggle with the vocalization of the next thought unit (Goffman 1981: 293). It could therefore help the listener to understand that they can expect more information to follow whilst simultaneously preventing them from taking over the turn before the previous speaker has given the floor up voluntarily.

8. Discussion

Reviewing the initial hypotheses it is surprising how much the results deviate from the expected outcome.

8.1 Review Hypothesis I

I. uh/um hesitation markers will occur slightly less frequently within the deceptive statements than within the truthful statements.

Not only do the uh/um hesitation markers occur less frequently; they do so by a considerable margin. In deceptive speech this specific uh/um hesitation occurs at only 43.2% of the rate that it would occur in truthful speech. This goes completely against the findings by both DePaulo et al. (2003) and Vrij et al. (2008), who had concluded that either no difference in hesitation usage occurs in either speech scenario (Vrij 2008: 94, 124) or that when it does, said difference could not be deemed statistically significant (DePaulo 2003: 97). The findings of this paper however, are more in line with those of Villar and Arciuli (2012). In their study of the speech pattern of a convicted murderer they had found that within truthful speech the uh/um hesitation markers make up 5.22% (Speech Scenario I) and 3.93% (Speech Scenario II) of all words employed[5] (Villar and Arciuli 2012: 91). In the deceptive statements uttered in speech scenario I & II the usage of the same hesitation markers sank to 1.11% and 0.12% respectively. The combined averages of all speech scenarios in this paper divided into the truthful/deceptive categories results in a 4.9% distribution of uh/um in truthful speech and in only 1.7% for deceptive speech.

These extreme and often paradoxical results of the different studies lead to the question as to why such differences could occur to begin with.
When looking at the individual studies incorporated into the meta analysis by both DePaulo et al (2003) and Vrij et al (2008), the variety of sources and methodology employed by each individual study becomes apparent, leading one to argue that the initial positions of each study are too different to compare adequately in one meta analysis without employing measures to counteract said differences. One example of this would be that most studies within the meta analysis do not unanimously focus on the same hesitation markers in their research. Dulaney

[5] Including uh/um as words

(1982)[6] for example does not specifically state exactly which hesitations are included within their analysis, as they had only mentioned "ah" and "uh" as examples of these (Dulaney 1982: 78). Furthermore Dulaney (1982) counted these hesitations towards overall speech disfluency ratio instead of viewing them by themselves (Dulaney 1982: 79). Contrary to that Miller et al. (1983)[7] included at least "uh", "ah" "um" and "mm" in their analysis (Miller et al. 1983: 108). While Miller and Stiff distinguish these speech errors with other errors such as repetitions they also do not look at the hesitation markers individually.
Vrij et al. (2008) do in fact comment on this problem of "inadequate scoring systems" (Vrij et al. 2008: 62) by stating:

> Although not yet investigated, a similar pattern may emerge if we examine in more detail another vocal category, speech hesitations ("ah", "um", "er", "uh", and "hmmm"). Smith and Clark (1993) found that "um" indicates higher cognitive load than "uh". Therefore examining "ums" and "uhs" separately may be indicative of lying. Perhaps liars include more "ums" in their statements than truth tellers, due to enhanced cognitive load, but do not differ in their use of "uhs".
>
> (Vrij et al. 2008: 62)

Another issue is the question of deception elicited within lab-setting versus nature-setting and the position of the deceptive individual. As before mentioned Ekman (1985/1992) had noted difficulties with recreating deception in a lab setting and therefore had to re-conduct his experiment to increase the stakes of his second study (Ekman 1985/1992: 54-56). In Ekman's (1985/1992) case subjects of the first flawed study did not engage in their lie-telling process enough, which resulted in Ekman deeming the results unusable since they were unrepresentative of the actuality of deception (Ekman 1985/1992: 54). Not all studies in the meta analysis of both DePaulo et al. (2003) and Vrij et al. (2008) belong exclusively to either category of lab-generated or naturally occurring as well. A study conducted by Bond et al (2004)[8] for example, used undergraduate students who upon participation were promised extra credits, as the source for their data (Bond et al. 2004: 30). Even for an experiment conducted in a lab-setting, these prove to be considerably low stakes for the participants, as failing to avoid deception detection is neither rewarded nor punished in this study. Over the course of the experiment, Bond et al. (2004) concluded that hesitation occurs more frequently

[6] Part of DePaulo et al. (2003) meta analysis
[7] Also part of DePaulo et al. (2003) meta analysis
[8] Part of Vrij et al. (2008) meta analysis

within deceptive speech than in truthful speech. On the other hand Davis et al. (2005)[9] conducted their experiment on testimonies given in real life scenarios in which 28 criminal suspects[10] gave their testimony to an assistant district attorney (Davis et al. 2005: 688). As the charges in all cases were rather intense (murder, homicide, assault, etc.), there was ample motivation for the candidates to avoid apprehension. Contrary to Bond et al. (2004) Davis et al. (2005) concluded hesitation to occur less frequently in deceptive speech than in truthful speech. Another issue that comes with this is that generally the participants in the lab-generated deception had little time to prepare their lies ahead as they were only informed of the procedure of the study upon participation. Contrary to that the subjects of the naturally occurring acts of deception all had time to prepare their story ahead of time as in each instant a significant amount of time had passed between the crime and the interview.

Galasinksi (2019) has commented on this exact problem stating:

> And so, this is why most deception research is often based either on statements or on vignettes whose origin is the imagination of the researcher, or, as in social psychology, on researchers' instructions for the participants to lie or deceive. But then it is possible to argue that such research is based on constructions of deception rather than deception itself. Researchers make up scenarios they imagine reflect what happens in reality, and at the same time, research participants play liars and their actions are 'as if' they were lying. Once could argue that such actions do not have to reflect actual communicative reality, where there might be high stakes associated both with deception and with being found out.
>
> (Galasinksi 2019: 521)

In Elizabeth Holmes' case we find an act of deception that is both highly motivated and meticulously prepared for. All research that used source material produced under similar circumstances (e.g. Villar and Arciuli 2012, Davis et al 2005, Vrij and Mann 2001, Vrij and Mann 2002) had also found hesitation to decrease during deceptive speech production. This therefore seems to suggest that preparedness and motivation to succeed play a huge role in hesitation production.

[9] Also part of Vrij et al. (2008) meta analysis
[10] Each suspect was later convicted of the respective crime in question

Leaving possible issues with the methodology of the meta analysis behind, there is a different set of concerns. One reason is that a person's speech pattern could simply be too individually distinct as to identify cues for deception that are universally applicable across the board. Researches have often noted that it seems impossible to accurately pinpoint one cue as indicator for deceptiveness and have argued a general change in behavior and speech pattern that is individually identified for the person in question to be more reliable (Ekman et al. 1991: 132). Wieling et al (2016) had found that uh/um hesitation marker usage across Germanic languages differs highly dependent on the speaker's sociolinguistic background including factors such as age and gender (Wieling et al. 2016: 15). When it comes to the production of um over uh hesitation their study had shown that younger speakers tend to produce ums significantly more than an older speaker would (Wieling et al 2016: 15). Furthermore, they had found that women produce more ums than male speakers as well (Wieling et al 2016: 16). However overall, Tottie (2011) had found male speakers, older speaker and speakers over higher educational backgrounds to hesitate more (Tottie 2011: 192). Comparison between speakers of different socioeconomic background, gender and age therefore seems to be inherently skewed towards delivering contradictory results.

Lastly, there is a possibility that there is actually a decrease in cognitive load during deception which would correlate with the decrease of hesitation phenomena. Van Bockstaele et al. (2012) had found evidence suggesting lying becomes less cognitively demanding when practiced (Van Bockstaele et al. 2012: 6). Since Holmes has been participating in the same fraudulent scheme for an extended period of time (approximately 15 years) she is likely to have repeated her lies numerous times in various high stakes scenarios (namely meetings and interviews). She would therefore be well versed in the production of said lies and likely has certain deceptive answers to often asked question memorized so to not run into the issue of having to make up a lie spontaneously on the spot and possibly incriminating oneself due to contradictions of the narrative.

8.2 Review Hypothesis II

II. Pauses will be of longer duration during Holmes' deceptive utterances compared to her truthful ones. Pauses will however not be more frequent within deceptive utterances than within truthful ones.

Contrary to the hypothesis, the results of this experiment show that pause length actually decreases in deceptive speech. Whilst the average pause length in the truthful speech scenarios I & II is 604ms and 543ms (combined average 556ms) respectively, the average pause length in the deceptive speech scenarios III, IV, V and VI drops to 460ms, 410.5ms, 403ms and 412ms (combined average 430ms). These findings go against both the results of the meta analysis by DePaulo et al. (2003) as well as Vrij et al. (2008). Furthermore, in the case of Elizabeth Holmes' the frequency of pauses indeed varies between truth-telling vs. lie-telling. In truthful speech the average lies at 24.7 pauses per minute, whereas in deceptive speech that average lies at 13.8 pauses per minute. So not only are pauses shorter in deceptive speech but also significantly less frequent. Whilst there were paradoxical findings in the individual studies of Vrij et al. (2008) when it came to hesitation markers, none of the studies had supported a decrease in pause duration during deception (Listed in Vrij et al. 2008: 90-94). This again could be explained by the before mentioned findings of Van Bockstaele et al. (2012), making pauses also less likely to occur in scenarios of well-practiced lies. However, a reason outside of potential cognitive factors could be the nature of the utterances themselves. Rühlemann et al. (2013) had found evidence to suggest pauses, as well as er/uh and erm/um hesitation markers occur more frequently in conversational narrative, aka story-telling as opposed to speech preceding and succeeding narration (Rühlemann et al. 2013: 68). When looking at Holmes' interactions one can see that the statements given in the truthful speech segments are much more of a narrative nature as she is mostly retelling stories from her time spent at high school and at Stanford. Contrary to that the statements given in the deceptive speech segments are much more of an explanatory nature, serving the purpose of detailing the function of and proceedings in Theranos and their systems. Because of this there could be an argument made that not only the act of lying versus truth-telling plays a role in the distribution of both pauses and uh/um hesitation markers but also the purpose of the speech acts.

7.3 Review Hypothesis III

Hypothesis 3: uh/um hesitation markers will not occur arbitrarily throughout the utterances, but there are points in sentence structure where they are statistically more likely to occur. These points are presumably the phrase boundaries.

The analysis has shown that in Holmes' speech there is the following distribution of uh/um hesitation markers: 20.8% of hesitation markers are found at the immediate beginning of a new sentence. 54.7% of markers are found in and immediately around the conjunctive element of compound sentences or complex sentences. The remaining 24.5% are found within the sentence, almost exclusively at the phrase boundaries. There is only one instance out of the 53 hesitations where it can be argued to have occurred at the word boundary inside of one Verb Phrase, yet even this is dependent on the researcher's approach of non-finite Verb Phrase position within constituency structure. Not a single hesitations was found to be dividing a word/compound word, i.e. "high uh school". Since the immediate beginning of a new sentence and the conjunction word itself represent the beginning/end of a new phrase, they too represent phrase boundaries. Therefore, an overwhelming majority of uh/um hesitation markers occur at these boundaries, supporting Maclay and Osgood's (1959) findings. However, Maclay and Osgood's (1959) other findings, those of hesitation preceding predominantly lexical words over function words is unsupported by the results of this paper, which had found function words to be much more prominent collocations of hesitation markers. However, it is important to note here that Maclay and Osgood (1959) had followed the definition Fries (1952) gave for lexical words and function words. They had noted here that in his categorization Fries (1952) had made a distinction between personal and possessive pronouns and had put personal pronouns into the same category as nouns (Maclay and Osgood 1959: 27). Since a substantial amount of collocations are subject personal pronouns in both our analysis as well as Kjellmer (2003) and Rühlemann (2010), Fries' (1952) inclusion of these within lexical words explains why the results in Maclay and Osgood's (1959) research differ so much when we traditionally expect hesitation to occur more frequently preceding function words.

Out of the 53 uh/um hesitation markers in this study 39 (75%) are succeeded by function words, whereas only 12 (22.6%) are followed up by lexical words. The remaining 2 are followed by other interjections such as "well" or "you know", so they were removed from the

count. A majority of these collocations are conjunctions (34%) and pronouns (30%). In the L1 location of uh/um hesitation markers 18 collocations are conjunctions whereas in the R1 location 19 collocations are conjunctions and 15 are pronouns. In their study Rühlemann et al. (2013) found that pauses, as well as er/erm[11] hesitation markers are most often succeeded by the pronoun "I" (Rühlemann et al. 2013: 71-70). This collocation occurs not only first succeeding word (R1 location), but also at the second (R2 location) and third word (R3 location), albeit with lowered chance (Rühlemann et al. 2013: 71). It was further noticed that pauses are often collocates of the conjunction "and", especially in narrative conversation (Rühlemann et al. 2013: 70).

So what does this mean for lexical status of uh/um hesitations?

As stated in 4.3 there is an ongoing debate on the status and function of hesitation markers within language and discourse. De Leeuw (2007) had described the two opposing theories by the terms "symptom hypothesis" and "signal hypothesis" respectively (De Leeuw 2007: 86). Both sides of the theory have given ample examples of what they believe to be indicative of their side of the hypothesis. For example, one of the main arguments of Clark and Fox Tree (2002) for uh/um hesitation markers being interjections was the emergence of a usage pattern within sentence structure that mimicked the behavior of traditional interjections in syntax (Clark and Fox Tree 2002: 103-104). Whilst the analysis done within this paper supports the notion of uh/um hesitation not occurring arbitrarily, but favoring certain positions within sentence structure this does not automatically mean that this is due to their supposed lexical importance. Clark and Fox Tree (2002) defined uh/um hesitation marker "syntax" as the following:

> Interjections, by definition, do not take part in syntactic constructions, although they get their meaning in part from their placement within such constructions. When George says "Sunday . the twenty-fifth, - sorry twenty-fourth […]", for example, he uses "sorry" to point backward to a recent offense (the error "twenty-fifth") and to point forward to a fix for the offense (the repair "twenty-fourth"). Uh and um are no

[11] er/erm = uh/um

> different. Speakers use them to pinpoint the initiation of an expected delay and to point forward to future speech.
>
> Clark, Fox Tree (2002: 104)

This comparison of uh/um to a secondary interjection seems a bit far-fetched. The term "sorry" is a very deliberate speech construction that demands a number of specific phonemes to be vocalized in a specific order (ˈsɒrɪ/ˈsɑːri). A speaker arguably does not construct interjections like "sorry", "I mean" or "oh my god" accidentally, but chooses specific words which are semantically appropriate for the context of the conversation over a number of other primary and secondary interjections. It would therefore be more appropriate to compare uh/um to other non-words aka primary interjections such as "ah" or "mhm".
Another argument this paper would like to address is one by Arciuli and Villar (2012) who had stated that:

> The current study[12] demonstrates in a real-word forensic context the discriminative ability of the use of "um." The results suggest that in an attempt to successfully deceive, humans may strategically alter their linguistic behavior. Furthermore, these findings suggest that a (re)conceptualization of "um" as a lexical term that is under the control of the speaker, as opposed to an unplanned speech error or filler, warrants further investigation.
>
> Arciuli and Villar 2012: 92.

While there are indeed several qualities of (uh/)um hesitation markers that suggest these to be interjections as we will see in the following paragraph, there is simply a lack of evidence to suggest that the supposed "suppression" of um hesitation markers is intentional and not because of a variety of other factors such as preparedness or increased focus on speech production. So while an ability to suppress the exclamation of uh/um could indeed be indicative of uh/um being actively employed by the speaker as opposed to unconsciously uttered, Arciuli and Villar's (2012) work does not offer enough evidence to support the idea that the lack of hesitation markers in deceptive speech is due to speaker control. This paper therefore wishes to caution as not to identify the analysis of Holmes's speech and the

[12] Um usage of a convicted murderer in both truthful and deceptive statements

correlating results as undisputable evidence for the lexical position of uh/um hesitation markers as interjections.

However, on the other side of the hypothesis there is indeed an argument to be made that uh/um could indeed be primary interjections. Primary interjections are words that cannot be used outside of their specific function as an interjection, as opposed to secondary interjections which have multiple applications due to their notional semantics (Ameka 1992: 105). Primary interjections therefore include terms such as "gee", "ugh" or "phew", whereas secondary interjections use "proper" words such as "heavens", "holy hell" or "thank God". As uh/um certainly do not serve a function across several word categories but are confined to their role as interjection, they are to fall within the category of primary interjection. As interjections refer to the speaker's mental state or mental act, we can further define uh/um hesitation markers in their expressive function (Wierzbicka 1992: 165). The three categories to choose from here are: emotive function, volitive function or cognitive function. Wierzbicka (1992) describes these as such:

1) emotive ones (those which have in their meaning the component 'I feel something');

2) volitive ones (those which have in their meaning the component 'I want something' and which do not have the component 'I feel something'; e.g. the English Sh! […])

3) cognitive ones (those which have in their meaning the component 'I think something' or 'I know something' and which have neither the emotive component 'I feel something' nor the volitive component 'I want something'; e.g. the English Aha! 'I understand'

Wierzbicka 1992: 165

As before mentioned, Clark and Fox Tree (2000) have argued that the usage of uh/um serves as an indicator to let the listener know that previous or upcoming speech disruption has been detected and efforts are being made to correct these (Clark and Fox Tree 2000: 104). This would indicate uh/um hesitation markers falling into the category of having cognitive function, as they fulfill the demand for "I know something" that being "I am aware that speech disruption is occurring/about to occur and I am actively working to fix it". However,

whilst this categorization of uh/um hesitation markers as interjections seems plausible there is currently not enough reliable evidence in previous research to indicate the speaker produces uh/um solely with the intent to signal errors in speech production and not just as a reactionary symptom.

Corley and Stewart (2008) summarized it best by stating:

> In the absence of such evidence, the picture remains murky. There is no conclusive evidence that fillers are words. What can be surmised at present is that, like a facial gesture or a tone of voice, hesitation disfluencies like um provide information to the listener. The information is something like "pay attention, the speaker's in trouble and the next part of the message might not be what you predicted"
>
> (Corley and Stewart 2008: 14)

9. Conclusion

In this analysis changes in Elizabeth Holmes's speech pattern regarding her usage of uh/um hesitation markers and pauses during truthful and deceptive speech acts have been analyzed in order to establish if said linguistic cues could potentially be viewed as indicators for deception taking place. Furthermore, this paper tried to investigate claims of uh/um hesitation markers having the same lexical properties as interjections by examining if distribution of these within sentence structure could uncover more evidence for speaker control. Though further research on the topic is encouraged, the current results suggest uh/um hesitation to occur far less frequently in deceptive speech than truthful speech. Holmes' changes in speech pattern also show a decrease in both pause length and pause frequency in deceptive speech. It is important to note here that these results are not exemplary for every kind of lie across the board, i.e. white lies, omissions, evocations etc., but for those that occur in high stakes scenarios in which the speaker has prepared the act of deception ahead of time. Earlier research has furthermore shown that the conditions of the speech scenario are vital in how one needs to evaluate the changes in speech pattern, especially when it comes to the question of lab-generated deception and naturally occurring deception. In addition to that, the sociolinguistic background of a speaker has further implications for their respective hesitation marker usage, though majority of speakers will show a change in the frequency of uh/um when under cognitive load. This paper has furthermore found results in support of previous research

concerning uh/um hesitation marker “syntax”. The results presented here show that uh/um hesitation occurs predominantly before and between clauses. Because of this uh/um is also overwhelmingly found at the phrase boundaries predominantly initiating Noun Phrases. This leads to the collocations of uh/um hesitation being primarily conjunctions such as “and” in the L1 location and both conjunctions and personal pronouns such as “I” in the R1 location. However, though this shows hesitation to not occur randomly throughout sentence structure it does not serve as reliable evidence to indicate uh/um is being deliberately employed by a speaker to warn the listener of oncoming or ongoing speech production errors.

Future research on the linguistic cues of deception will have to take a more precise approach to the selection of their source pool as to not run the issue of inadequately comparing contrasting origins of lie-telling to one another. Lying is a highly subjective experience and dependent on a number of factors such as sociolinguistic background, preparedness and intelligence of the speaker, all of which have to be considered in the analysis. This paper would therefore suggest focusing on naturally occurring lies which contents and origins are comparable to one another. Furthermore, proponents of the signal hypothesis over the symptom hypothesis of uh/um hesitation markers could profit from more reliable evidence. Current research methodologies, while helpful with the localization of hesitation markers in syntax, do not deliver results which can reliably prove the speaker’s control over the production of these, but only present results which are merely suggestive. In the future, research could benefit from taking a closer look at the speaker and listener interaction to find how turn-taking and general comprehension is influenced by a lack or an increase of uh/um hesitation.

Words: 14.882

Symbols: 77.501

Appendix A

The transcription of the following interviews is based on DuBois et al (1993). However, as not to make the transcript too illegible, due to an unnecessary amount of markings, only the aspects relevant to this analysis will be transcribed. This includes pauses, hesitation, stutters and pre-emptive cut-offs of sentences. The commas represent what is deemed to be intonation unit boundaries however, since this paper did not use a program to analyze for these, these are based on hearing and are subject to potential human error.
The audio files corresponding audio files are labelled Interview I-VI and can be found within the E-Mail sent to the Prüfungsamt der Philiologie.
Note: The questions asked by the interviewer are merely paraphrased. They also do not count towards word and time count.

Truthful Interviews

Interview I: Elizabeth Holmes – Progressing as a Teenage Entrepreneur
Source: Stanford Technology Ventures Program
Year: 2009

Q: What was the process of dropping out of Stanford University at age 19 like and what advice did Holmes get?

Statement a:
um …(1.52) interestingly it was, a fairly binary decision. …(0.39) um I got a point, where, I-- …(0.4) I actually originally did not ..(0.24) intend ..(0.21) to drop out of Stanford but, (0.24) I wasn't going to any classes and I was spending all of my time talking to VCs, and so then, logistically it just seemed like a waste of money, because I was, you know, taking twenty units and …(0.68) I wasn't showing up, so um …(0.7) so, so originally the concept was well, …(0.5) I …(0.41) take a, a leave of absence …(0.76) and then …(0.32) uh … (0.87) it became really clear that, you know, I-I was at a point where …(0.61) another few …(0.65) classes …(0.59) in chemical engineering, was not necessary, for what I wanted to do …(0.49) and so, um …(0.92) so then …(0.87) I just made a decision that I was gonna figure out how to make it work …(0.76) and, uh …(1.35) (INAUDIBLE) and I, I actually left, I, I …(0.8) I ended up living in the basement, of a bunch of …(0.31) of Stanford grads, and started …(0.42) working from the basement of their house, and, uh …(0.72) and then finally

convinced someone to ..(0.26) to give us-- uh originally it was a bridge loan, and uh then …(0.72) it went through the process of ..(0.21) of starting to build the company.

Words: 194
Um: 4
Uh: 5
Pauses: 28
Time: 1 min 28 sec

Interview II: Jonathan Zittrain in Conversation with Elizabeth Holmes
Source: Bergman Klein Center for Internet and Society
Year: 2015

Q: What would a typical day look like in 9th grade for Holmes?

Statement a:
Yep, um …(0.3) you know, I …(0.72) I was, uh …(0.95) a young girl who, loved …(0.47) science, and math, and engineering. I had um …(0.9) family that …(0.74) um …(0.67) brought me up in such a way in which …(0.68) when my brother was getting …(0.48) Lego sets and construction sets and …(0.57) my friends were getting Barbies, I was getting, Lego sets and construction sets, and I think that …(0.55) was a really wonderful and special thing, um …(0.83) and um ..(0.25) I ..(0.27) I, er, was ..(0.25) very normal …(0.45) in the context of …(0.76) working to pursue something that ..(0.29) that I loved.

Q: What would a normal weekend look like for Holmes in high school?

Statement b:
Um well I started uh a company when I was in uh …(0.55) uh high school um that was …(0.88) um ..(0.28) distributing C++ programming compilers to …(0.33) um …(0.35) universities in China because I had …(0.79) studied Mandarin.

Q: What was the early process of starting Theranos like?

Statement c:

Um I …(0.329) I found what I loved …(0.43) and …(0.85) I …(0.85) um …(0.59) I- I'd always …(1.33) believed that …(0.36) the purpose of building something (0.26) building a company, is to …(0.65) make a difference in the world ..(0.29) and …(0.49) um …(0.86) I got to the point in which I, I found what I felt like I was born to do …(0.41).

Q: Holmes is asked to go into more detail about statement c and explain why she decided to focus on blood-testing

Statement d:

I …(0.41) er so I (INAUDIBLE) when we were talking about my, my family in …(0.43) disaster relief I'd ..(0.21) I'd spent a lot of time …(0.57) thinking about …(0.24) what does it mean to, make a difference …(0.32) and …(0.76) um ..(0.28) to me ..(026) what I, resonated with …(0.45) very very deeply was …(0.71) the experience of …(0.72) watching …(0.43) someone I loved …(0.7) get diagnosed with cancer and …(0.74) um …(0.57) understanding what I meant ..(0.23) to …(0.93) found out …(0.49) too late …(0.31) in the disease progression process in a case where …(0.78) if we'd known ..(0.28) earlier …(0.82) there was everything …(0.32) that could've been done, um ..(0.24) and ...(0.74) really believing fundamentally that if there was one thing that I could do with my life …(0.71) um …(0.8) that …(1.02) that I wanted to pursue, it would be trying to change that …(1.29) and so I got interested in, blood and labtesting because …(0.52) it's a tool (0.27) for being able to ..(1.17) help create a system ..(0.29) in which early detection, becomes a reality.

Q: How did Holmes gain the knowledge necessary for starting Theranos?

Statement e:

Wh-I ..(0.26) I had …(0.42) had the opportunity at Stanford. I was, I was an undergrad. um I'd …(0.67) sort of convinced my way, into working in one of the PhD programs and …(0.58) um …(0.35) and had ..(0.21) wonderful …(0.7) exposure and training ..(0.28) from that …(0.61) in the short time that ..(0.28) I was there um ..(0.28) but …(0.64) I- I think …(0.65) e- we reached a point where-- …(0.46) and I reached a point, where I knew …(0.86) what it

was that I was trying to solve for …(0.31) and decided that I wanted to go spend all of my time on that.

Q: Did Holmes have a "eureka moment" in the process of starting the company that lead her to this idea?

Statement f:

I-it was the decision that …(0.45) that calling, and that mission …(0.67) was what I wanted to do with my life ..(0.25) and, I was at a point in which …(0.55) I was spending …(0.59) 24 hours a day, on it, and my parents were spending, all this money, on all these classes, that I wasn't going to, and um …(0.59) that didn't make any sense, so, um …(0.8) so I decided to go spend ..(0.24) all my time on it, and um …(0.33) I'd ..(0.24) I'd found what I loved so--

Q: What was Holmes' parents reaction to her wanting to drop out to start a company?

Statement g:

Ah, they, they were amazing, I mean they …(0.69) they um, they'd saved up all their lives, for me to be able to go to Stanford, and um …(0.79) and the money, that ..(0.23) that they saved up would have let …(0.39) them retire, and they let me take that money, and put it into starting that company.

Words: 515
Uh: 5
Um: 23
Pauses: 101
Time: 3 min 45 sec

Deceptive Interviews

Interview III: Theranos' Elizabeth Holmes on the Lifeblood of the Internet – FULL CONVERSATION

Source: Vanity Fair

Year: 2015

Q: Holmes is asked to explain the technology of the nanotainer[13].

Statement a:

um …(0.3) so we, we invested …(0.76) a-a lot, in being able to, re-develop every test, or every assay that's run in a traditional laboratory …(0.34) to be able to run those tests on …(0.55) uh**,** tiny drops of blood, and so …(0.31) for example this summer our, first …(0.97) FDA clearance was on a test that, uh, traditionally required, uh, a huge amount of blood, ah, through these traditional, vacutainer tubes, and …(0.52) we redeveloped, the chemistry, we redeveloped the hardware, we redeveloped the software, we redeveloped the collection tubes …(0.3) uh**,** to be able to make it possible …(0.41) to do, these tiny sample based testing and we've now …(0.64) submitted 130 …(0.42) of these, in, pre-submissions to the FDA to take through, this regulatory process …(0.52) and, it really meant, redeveloping the whole infrastructure because …(0.55) y-y-you just can't do this, using the traditional framework and …(0.45) part of that is because, we're trying to solve a different problem, which is access to …(0.56) actionable information, a-at the place and at the time …(0.35) that the patient sees the physician, or that the patient shows up, to be able to get access to their own health information …(0.54) um, I-I think ..(0.25) I mean we are, we're standing on the shoulders of giants in in a huge way in the context of ..(0.29) the developments …(0.41) that we've seen in the technology sector, and the ability to apply those ..(0.4) to, some of these core fundamental areas, in ..(0.28) lab medicine, and in, other areas in healthcare.

Words: 230
Um: 2
Uh: 4
Pause: 22
Time: 1 min 30 sec

[13] Vial used in the Edison machine to store blood

Interview IV: Theranos CEO Elizabeth Holmes: Firing Back At Doubters | Mad Money | CNBC

Source: CNBC

Year: 2015

Q: How many blood-test does Theranos currently do?

Statement a:

Yes, so we had, communicated to the Wallstreet Journal, that we have submitted over 130 pre-submissions to FDA, with tests running on our proprietary devices …(0.49) um and have been taking those through the FDA submission process …(0.44) every test, that we offer in our laboratory, can run, on our proprietary devices (0.36), we bring tests up, on our proprietary devices, based on the frequency with which they're run …(0.55). So at any given point in time …(0.38) uh, we're running the tests that are most commonly ordered, um, but we've also …(0.41) done a lot of work, as part of this commitment that we've made and, it's been very controversial …(0.38) that we've actually become the first company advocating, for FDA regulation ..(0.25) of lab developed tests, and as part of that …(0.54) we, have, said, that we think every lab developed test really should go, through the FDA …(0.51) submission process, and so we've been consistent with it, and in fact …(0.42) we even just recently took our nanotainers …(0.34) through the FDA, clearance process, and sent submissions in, for those …(0.39) and as part of that process …(0.43) we're not even using …(0.53) our nanotainers except for the FDA cleared assays, um, so that every single thing that runs on our platform …(0.31) is getting to the point that it's gonna be FDA cleared.

Q: Holmes is asked on her opinion on comparing Theranos lab data with that of QuestLab[14].

Statement b:

W-we've already done it …(0.3) We've already done it. Absolutely. And it's actually even published, in our FDA decisions summary from this summer …(0.43) from, a 900 patient study, where we got FDA clearance of the exact system that the Journal is questioning …(0.38) and demonstrated, venous versus finger-stick, across a huge number of patients. It was 889, I think, for that test …(0.37) and we've done that, over and over again for every single test.

[14] QuestLab is a traditional laboratory and one of Theranos' biggest competitors

Words: 280
Um: 3
Uh: 1
Pause: 20
Time: 1 min 34 sec

Interview V: Elizabeth Holmes defends Theranos amid media scrutiny at Fortune's Global Forum | Fortune
Source: Fortune Magazine
Year: 2015

Q: How many tests is Theranos currently doing via the finger-prick method?

Statement a:
Yeah, so right now, just because of this FDA transition, (INTERRUPTION) we're only doing one ..(0.24) right …(0.3) but that doesn't mean we don't have the technology to do them, or that we haven't done them in the past.

Q: How confident is Holmes that Theranos will get FDA approval for all 200 standard blood-analysis tests?

Statement b:
Well we've …(0.42) er, so there's multiple parts to this, one is …(0.57) the validation of those tests under the lab framework, the other one is the FDA submission, and this is an area of evolving policy, that we're working to create a leadership position in. …(0.48) We've never, talked, publicly about, what the status of what we're doing with FDA is …(0.39) because we just hadn't felt, that was appropriate, um, until we got into all of this, recent press communications …(0.53) but, um, we're incredibly confident in the data, that we've submitted to the agency, we've worked on it for two years …(0.39) and, we've met the standard of comparing our tubes to, vacutainers, which are he big tubes that come from the arm …(0.32) and we're hopeful, uh, that we'll see we'll see those-- (INTERRUPTION)

Q: Does Holmes think she should've been more transparent with Theranos' data?

Statement c:

I-I think right now, we wanna try to lead in transparency, I mean, I made a call that, if we would be the first lab, with the confidence to submit thi- our test to FDA, that that would be the way to get the data out …(0.4) but there is no reason we can't do peer-review, and there is no reason we can't publish other stats, and we're gonna do that.

Words: 230
Um: 2
Uh: 1
Pauses: 10
Time: 1 min 6 sec

Interview VI: Getting Blood Work Done with Elizabeth Holmes of Theranos | Disrupt SF 2014

Source: TechCrunch

Year: 2014

Q: Holmes is asked to comment on what happens once the blood is drawn and how the technology of the Edison machine works

Statement a:

Sure, exactly, so from the time you're …(0.44) in the wellness-center …(0.49) once, the price points have been identified, for, someone, so that they know-- Even if they have a deductible-- How much it's gonna cost them, and they can ..(0.24) decide, up front, whether they wanna purchase that service, as opposed to …(0.48) getting a bill, in the mail, you know, three months later, for, some unknown amount-- …(0.55) um ..(0.21) The …(0.32) little nanotainer that you saw, has a barcode on, it which tracks ..(0.23) the sample, and …(0.51) [when] the samples get to our labs, what we've done, is redevelop …(0.54) all of the chemistry …(0.34) associated with running …(0.48) these tests, on traditional platforms, to make it possible to …(0.55) run ..(0.27) any laboratory test, from a tiny droplet of blood …(0.37) and so …(0.42) what will happen to that sample is, it will be run through …(0.56) those chemistries ..(0.25) and we've also built out …(0.78) novel, analytical systems, on which to run ..(0.28) the chemistries, because traditional instrumentation, requires,

much larger tubes of blood …(0.61) and then that data, will be, electronically sent …(0.53) to, the ordering physician, uh, and integrated into …(0.43) uh, their EMR systems, for example, electronically.

Words: 179
Um: 1
Uh: 2
Pause: 23
Time: 1 min 11 sec

Bibliography

Ameka, F. (1992). Interjections: The universal yet neglected part of speech. *Journal of Pragmatics, 18*(2-3), 101-118. doi:10.1016/0378-2166(92)90048-g

Anthony, L. (2019). AntConc (Version 3.5.8) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software

USA: CLINTON DENIES HAVING AN AFFAIR WITH MONICA LEWINSKY UPDATE | AP Archive. (1998, January 26). Aparchive.Com. http://www.aparchive.com/metadata/youtube/ac91ce9759b801d480b6530eefe8c04d

Audacity(R): Free Audio Editor and Recorder (2.0.5). (2020). [Computer program]. Audacity Team. http://audacity.sourceforge.net/

Austin, J. L. (1962). *How to do things with words*. London, United Kingdom: Oxford University Press.

Bavelas, J. B., Black, A., Chovil, N., & Mullet, J. (1990). *Equivocal communication*. Newbury Park, CA: Sage Publications.

Bryant, E. (2008). Real Lies, White Lies, and Gray Lies: Towards a Typology of Deception. *Kaleidoscope: A Graduate Journal of Qualitative Communication Research*. 7, 23-48. https://digitalcommons.trinity.edu/hct_faculty/6/#:~:text=White%20lies%20were%20perceived%20as,as%20justifiable%20given%20the%20circumstance.

Bloomquist, J. (2010). Lying, cheating, and stealing: A study of categorical misdeeds. *Journal of Pragmatics, 42*(6), 1595-1605. doi:10.1016/j.pragma.2009.11.008

Van Bockstaele, B., Verschuere, B., Moens, T., Suchotzki, K., Debey, E., & Spruyt, A. (2012). Learning to Lie: Effects of Practice on the Cognitive Cost of Lying. *Frontiers in Psychology, 3*. doi:10.3389/fpsyg.2012.00526

Bond, C. F., Thomas, B. J., & Paulson, R. M. (2004). Maintaining lies: The multiple-audience problem. *Journal of Experimental Social Psychology, 40*(1), 29-40. doi:10.1016/s0022-1031(03)00087-8

Boomer, D. S. (1965). Hesitation and Grammatical Encoding. *Language and Speech, 8*(3), 148-158. doi:10.1177/002383096500800302

Brennan, S., & Williams, M. (1995). The Feeling of Another′s Knowing: Prosody and Filled Pauses as Cues to Listeners about the Metacognitive States of Speakers. *Journal of Memory and Language, 34*(3), 383-398. doi:10.1006/jmla.1995.1017

Buller, D. B., Burgoon, J. K., Buslig, A. L., & Roiger, J. F. (1994). Interpersonal Deception VIII: Further Analysis of Nonverbal and Verbal Correlates of Equivocation from the Bavelas et al. (1990) Research. *Journal of Language and Social Psychology, 13*(4), 396-417. doi:10.1177/0261927x94134003

Cambridge Dictionary (2020a). LIE: Bedeutung im Cambridge Englisch Wörterbuch. Retrieved October 24, 2020, from https://dictionary.cambridge.org/de/worterbuch/englisch/lie

Cambridge Dictionary (2020b). DECEPTION: Bedeutung im Cambridge Englisch Wörterbuch. Retrieved October 24, 2020, from https://dictionary.cambridge.org/de/worterbuch/englisch/deception

Carson, T.L. (2019). Lying and Ethics. In: J. Meibauer, *Oxford Handbook of Lying* (469-482). Oxford University Press.

Christenfeld, N. (1995). Does it hurt to say um? *Journal of Nonverbal Behavior, 19*(3), 171-186. doi:10.1007/bf02175503

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. The M.I.T. Press. doi:10.21236/ad0616323

Clark, H. H. & J. E. Fox Tree (2002). Using uh and um in spontaneous speaking. *Cognition, 84*(1), 73-111. doi:10.1016/s0010-0277(02)00017-3

Corley, M. & Stewart (2008) Hesitation disfluencies in spontaneous speech: The meaning of um. *Language and Linguistics Compass*, 2 (4), 589-602. https://doi.org/10.1111/j.1749-818X.2008.00068.x

Davis, M., Markus, K. A., Walters, S. B., Vorus, N., & Connors, B. (2005). Behavioral Cues to Deception vs. Topic Incriminating Potential in Criminal Confessions. *Law and Human Behavior, 29*(6), 683-704. doi:10.1007/s10979-005-7370-z

Depaulo, B. M., Kashy, D. A., Kirkendol, S. E., Wyer, M. M., & Epstein, J. A. (1996). Lying in everyday life. *Journal of Personality and Social Psychology, 70*(5), 979-995. doi:10.1037/0022-3514.70.5.979

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. doi:10.1037/0033-2909.129.1.74

DuBois, John & Schuetze-Coburn, Stephan & Cumming, Susanna & Paolino, Danae. (1993). Outline of discourse transcription. *Lawrence Erlbaum Associates Publishers*. Hillsdale, NJ.

Dulaney, E. F. (1982). Changes In Language Behavior As A Function Of Veracity. *Human Communication Research, 9*(1), 75-82. doi:10.1111/j.1468-2958.1982.tb00684.x

Ekman, P. (1992). *Telling lies ; clues to the deceit in the marketplace, politics, and Marriage*. New York, NY: W.W. Norton & Company.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *PsycEXTRA Dataset*. doi:10.1037/e525532009-012

Ekman, P., O'sullivan, M., Friesen, W. V., & Scherer, K. R. (1991). Invited article: Face, voice, and body in detecting deceit. *Journal of Nonverbal Behavior, 15*(2), 125-135. doi:10.1007/bf00998267

Fallis, D. (2019). Lying and Omission. In: J. Meibauer, *Oxford Handbook of Lying* (183-192). Oxford University Press.

Fries, C. C. (1952). *The Structure of English*. Harcourt, Brace & World.

Galasinksi, D. (2019). Lying and Discourse Analysis. In: J. Meibauer, *Oxford Handbook of Lying* (517-528). Oxford University Press.

Goldman-Eisler , F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. New York, NY: Academic Press.

Grice, Paul. (1991). *Studies in the Way of Words*. Cambridge, MA. Harvard University Press.

Goffman, E. (1981). Radio talk. In E. Goffman (Ed.),Forms of talk(pp. 197–327). Philadelphia, PA: Universityof Pennsylvania Press.

Heldner, M., & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics, 38*(4), 555-568. doi:10.1016/j.wocn.2010.08.002

Hieke, A. E., Kowal, S., & O'connell, D. C. (1983). The Trouble with "Articulatory" Pauses. *Language and Speech, 26*(3), 203-214. doi:10.1177/002383098302600302

James, D. (1972). Some aspects of the syntax and semantics of interjections. Papers from the eighth regional meeting of the Chicago Linguistic Society (pp. 162–172). Chicago, IL: Chicago Linguistic Institute.

Kasl, S. V., & Mahl, G. F. (1965). Relationship of disturbances and hesitations in spontaneous speech to anxiety. *Journal of Personality and Social Psychology, 1*(5), 425-433. doi:10.1037/h0021918

Keltner, D., & Harker, L. A. (1998). Forms and functions of the nonverbal signal of shame. In P. Gilbert & B. Andrews (Eds.), *Interpersonal approaches to shame* (pp. 78–98). Oxford, England: Oxford University Press.

Kjellmer, G. (2003) Hesitation. In Defence of ER and ERM, *English Studies*, 84:2, 170-198, doi:10.1076/enst.84.2.170.14903

de Leeuw, Esther. (2007). Hesitation Markers in English, German, and Dutch. *Journal of Germanic Linguistics, 19*(02). doi:10.1017/s1470542707000049

Maclay, H., & Osgood, C. E. (1959). Hesitation Phenomena in Spontaneous English Speech. *WORD, 15*(1), 19-44. doi:10.1080/00437956.1959.11659682

Mann, S. (2019). Lying and Lie Detection. In: J. Meibauer, *Oxford Handbook of Lying* (408-419). Oxford University Press.

Miller, G. R., Deturck, M. A., & Kalbfleisch, P. J. (1983). Self-Monitoring, Rehearsal, And Deceptive Communication. *Human Communication Research, 10*(1), 97-117. doi:10.1111/j.1468-2958.1983.tb00006.x

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying Words: Predicting Deception from Linguistic Styles. *Personality and Social Psychology Bulletin, 29*(5), 665-675. doi:10.1177/0146167203029005010

Nicklaus & Stein. (2020). The Role of Linguistics in Veracity Evaluation. *Journal of Language and Law* 9, 23–47, doi:10.14762/jll.2020.023

Osgood, C. E., Sebeok, T. A. (Eds.), & Gardner, J. W., Carroll, J. B., Newmark, L. D., Ervin, S. M., Saporta, S., Greenberg, J. H., Walker, D. E., Jenkins, J. J., Wilson, K., & Lounsbury, F. G. (1954). Psycholinguistics: a survey of theory and research problems. *The Journal of Abnormal and Social Psychology, 49*(4, Pt.2), i–203. https://doi.org/10.1037/h0063655

Ortony, A. & Gupta, S. (2019). Lying and Deception. In: J. Meibauer, *Oxford Handbook of Lying* (149-169). Oxford University Press.

Quinting, G. (1971). Hesitation phenomena in adult aphasic and normal speech. Paris, *Mouton*.

Rühlemann, C., Bagoutdinov, A., & O'donnell, M. B. (2011). Windows on the mind. *International Journal of Corpus Linguistics Errors and Disfluencies in Spoken Corpora, 16*(2), 198-230. doi:10.1075/ijcl.16.2.03ruh

Salvetti, Franco (2014). Boulder Lies and Truth LDC2014T24. Philadelphia: Linguistic Data Consortium. https://catalog.ldc.upenn.edu/LDC2014T24

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. London: Cambridge University Press.

Stokke, A. (2019). Lying, Sincerity and Quality. In: J. Meibauer, *Oxford Handbook of Lying* (134-148). Oxford University Press.

Tenbrink, T. (2020). *Cognitive Discourse Analysis: An Introduction*. Cambridge University Press.

Tottie, G. (2011). Uh and Um as sociolinguistic markers in British English. *International Journal of Corpus Linguistics Errors and Disfluencies in Spoken Corpora, 16*(2), 173-197. doi:10.1075/ijcl.16.2.02tot

Villar, G., Arciuli, J., and Mallard, D. (2012). Use of “um” in the deceptive speech of a convicted murderer. Applied Psycholinguistics, 33(1), 83-95. doi:10.1017/S0142716411000117

Vrij, A., & Mann, S. (2001). Telling and detecting lies in a high-stake situation: The case of a convicted murderer. Applied Cognitive Psychology, 15(2), 187–203. https://doi.org/10.1002/1099-0720(200103/04)15:2<187::AID-ACP696>3.0.CO;2-A

Mann, Samantha & Vrij, Aldert & Bull, Ray. (2002). Suspects, Lies, and Videotape: AnAnalysis of Authentic High-Stake Liars. Law and human behaviour, 26, 365-76. 10.1023/A:1015332606792.

Vrij, Aldert & Edward, Katherine & Roberts, Kim & Bull, Ray. (2000). Detecting Deceit via Analysis of Verbal and Nonverbal Behavior. Journal of Nonverbal Behavior, 24, 239-263. 10.1023/A:1006610329284.

Vrij, A. (2008). Detecting lies and deceit: pitfalls and opportunities. West Sussex, Wiley.

Wells, R. S. (1947). Immediate constituents. Language, 23, 81–117. https://doi.org/10.2307/410382

Wieling, M., Grieve, J., Bouma, G., Fruehwald, J., Coleman, J., & Liberman, M. (2016). Variation and change in the use of hesitation markers in Germanic languages. Language Dynamics and Change, 6, (2), 199-234. https://doi.org/10.1163/22105832-00602001

Wierzbicka, A. (1992). The semantics of interjection: Journal of Pragmatics, 18, (Issues 2–3), 159-192, https://doi.org/10.1016/0378-2166(92)90050-L.

Zellner, B. (1994). Pauses and the temporal structure of speech. in E. Keller (Ed.) Fundamentals of speech synthesis and speech recognition. 41-62. Chichester: John Wiley.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In L. Berkowitz (Ed.), Advances in experimental social psychology, 14, 1–59. New York: Academic Press. https://doi.org/10.1016/S0065-2601(08)60369-X

Interviews

Stanford Technology Ventures Program. (2009). *Elizabeth Holmes-Progressing as a Teenage Entrepreneur* [Video]. Youtube. Last accessed 24.10.2020. https://www.youtube.com/watch?v=pIXAmjUFSso

Bergman Klein Center for Internet & Society. (2015). *Jonathan Zittrain in conversation with Elizabeth Holmes* [Video]. Youtube. Last accessed 24.10.2020. https://www.youtube.com/watch?v=uYRafsRu2NU&t

Vanity Fair. (2015). *Theranos's Elizabeth Holmes on the Lifeblood of the Internet – FULL CONVERSATION* [Video]. Youtube. Last accessed 24.10.2020. https://www.youtube.com/watch?v=hqQ3CMv_4XM&t

CNBC. (2015). *Theranos CEO Elizabeth Holmes: Firing Back At Doubters | Mad Money | CNBC* [Video]. Youtube. Last accessed 24.10.2020. https://www.youtube.com/watch?v=rGfaJZAdfNE

Fortune Magazine. (2015). *Elizabeth Holmes defends Theranos amid media scrutiny at Fortune's Global Forum | Fortune* [Video]. Youtube. Last accessed 24.10.2020. https://www.youtube.com/watch?v=A8qgmGtRMsY&t

TechCrunch. (2014). *Getting Blood Work Done with Elizabeth Holmes of Theranos | Disrupt SF 2014* [Video]. Youtube. Last accessed 24.10.2020. https://www.youtube.com/watch?v=AaTHtSy5le4&t

Extras (Not part of the paper, but a recommended read for any further context on Theranos):

Carreyrou, John (2018). *Bad Blood : Secrets and lies in a Silicon Valley startup* (First ed.). New York: Knopf.

Declaration of Authenticity

Hiermit versichere ich, dass ich die Arbeit selbständig angefertigt, außer den im Quellen- und Literaturverzeichnis sowie in den Anmerkungen genannten Hilfsmitteln keine weiteren benutzt und alle Stellen der Arbeit, die anderen Werken dem Wortlaut oder dem Sinn nach entnommen sind, unter Angabe der Quellen als Entlehnung kenntlich gemacht habe.

Sankt Augustin, den 25.10.2020

Datum, Unterschrift