

Puzzling Patterns: Exploring the Relationship Between Confidence and Filler Words in Teenage Speech

Varsha Bhargava¹ and Daria Kuscenko[#]

¹ Ridge High School, USA

[#]Advisor

ABSTRACT

The present study investigates whether an adolescent's confidence in their knowledge correlates to the frequency of filler words in their speech. High school participants (N = 31) were presented with two mazes, Maze A to evoke more confidence and obstacle-ridden Maze B to evoke less confidence. Participants submitted Likert Scales regarding their confidence in completing each maze, then completed the mazes. They were then recorded while discussing each maze. Finally, their speech was transcribed and annotated for filler words, marking distinctions between filled pauses and discourse markers. Using this data, a Spearman Rank Correlation Coefficient test and a paired t-test were conducted to test the strength of correlation between filler word frequency and confidence. The results of the Spearman Rank test implied negligible correlation between the variables, $r(29) = -.1, p = .184$. The paired t-test, using an accepted significance level of 0.05, indicated no difference in the means of filler word frequency data before confidence is decreased (M = 5.5, SD = 2.8) and filler word frequency after confidence is decreased (M = 5.7, SD = 3.1), $t(30) = .3, p = .740$. Additional Spearman Rank tests and paired t-tests results on filled pause and discourse marker usage before and after variation in confidence implied that an adolescent's confidence does not influence the type of the filler word they use. Due to sample size limitations, these conclusions cannot be generalized, so further research is needed, but this data lays the foundation for adaptations to school curriculums and societal perceptions of adolescents.

Introduction

As the negative stigma surrounding filler-word filled language prevails in modern conversation, it is important to evaluate the validity of this perception. A concurring stereotype perpetuated by modern media is the uncaring teenager whose favorite word is "like". This teenager's character is reflected and battered in the classroom by teachers who correlate filler word frequency and surety, then consequently perceive them as unconfident learners. Filler words have received an adverse reputation within the classroom, with many institutions publishing and teaching anti-filler word curriculums. Professors often expect students to eliminate filler words from their speech in order to effectively showcase their confidence in the subject they are speaking on. For example, Penn State University's SC200 course published this statement on their course website: "If someone is using a lot of filler words, they are not 100% confident on the points that they want to get across to their intended audience" in 2015 (Like like like like what?). This study aims to challenge this perception and investigate whether a true correlation exists between a teenager's filler word use and their level of confidence in a subject.

Literature Review

It is imperative to address the characteristics of filler words and confidence, as well as the existing conversations surrounding them, in order to answer the research inquiry for this study, which asks: how does level of confidence affect the frequency in which teenagers apply filler words in their speech?

Filler Words

Whereas filler words are often misunderstood or miscategorized in modern society and non-scholarly articles, a number of studies have shown that filler words are vocalizations unrelated to the speaker's topic that are used to fill pauses or hesitations in everyday speech (Gosy & Horvath, 2010). Additionally, many research findings show that filler words are more sophisticated than arbitrary vocalizations and often have underlying purposes, intentional or not. Filler words can fall under two categories: discourse markers and filled pauses (Bortfield et al., 2001). Recent experiments have demonstrated that rather than being superfluous, the specific types of filler words signal different intentions.

Discourse Markers

Accumulating research indicates that discourse markers are used to provide information about the speaker or situation, to indicate informality or soften a statement, to keep the audience engaged, to indicate casual approximation, or to signal that the speaker is contemplating their next thought (Fox Tree & Shrock, 2002). Generally, it has been noted that discourse markers are usually longer phrases, such as "I mean" and "like" (Laserna & Pennebaker, 2014; Bortfield et al., 2001).

Like all filler words, discourse markers are complex. Battistella's study investigated the single word "like" and all of the possible attributions to its usage; he explores many options for the one word, including attributing it to the speaker's stage presence, indication of approximation or limited knowledge on a topic, or marking indirect thought (Battistella, 2007). Battistella's 2007 study regarding a single filler word and all of its different characteristics suggests that there is more to be established about filler words and their interpretations.

Filled Pauses

Filled pauses, unlike discourse markers, are simple vocalizations such as "uh" and "um" that are used to fill pauses in speech (Alvar, et al. 2019). Researchers have raised questions regarding the meaning behind filled pauses, and several assumptions can be deduced from the current literature. Reasons such as expressing the process of decision-making and verbally marking pauses in thought without expressing emotion have been suggested as attributions to these vocalizations (Crible & Kosmala, 2021).

Overall, recent investigations interpret filler words as more complex and categorizable than before, allowing for more research to be conducted in this field of study.

Confidence

Confidence is an area to be explored and a number of hypotheses concerning the nature of its definition have been advanced by authors. The question "How is confidence defined?" has been a recurring motif in many studies in the field. Characteristically, the studies examined self-esteem, achievement, and experience when investigating the facets of confidence, however one highly critical factor was brought up frequently: self-efficacy (Chang & Yang, 2010; Kalra & Siribud, 2020).

Previous research indicated a significant relationship between confidence and perceived self-efficacy. In Kalra and Siribud's study, the researchers conducted interviews with their participants regarding their confidence levels in an activity and the feedback they received was that their confidence had a directly proportional relationship to their certainty in being able to complete the task (Kalra & Siribud, 2020). This approach will be mirrored and built upon in this present study.

Confidence in Relation to Filler Words and The Gap

A study concerned with the relationship between confidence levels and filler words is by Bortfield et al. in 2001, where they described their participants' levels of confidence to have waned as they began to doubt their answers to the experiment questions, which led them to fill their sentences with filler words. However, in this study, confidence was regarded as a measure of doubt, where the doubt and level of confidence had an inversely proportional relationship: the more doubt a participant had in their answer, the less confidence they had (Bortfield et al., 2001). While this definition worked for the context of Bortfield et al.'s study, the sense that self-efficacy is related to confidence, as established by Kalra and Siribud, is more pertinent to the goal of this study, which aims to solidify the relationship or lack of relationship between filler words and confidence as a measure of self-efficacy (Bortfield et al., 2001; Kalra & Siribud, 2020).

Additionally, regarding the sample populations in many of the foundational studies in the body of literature for this study, such as Bortfield et al. in 2001, Christenfeld, 1994, Crible and Kosmala, 2021, they were either performed with university students who were over 18 years old or the studies observed corpora in university contexts, and some of these studies were also focused on mainly non-native English speakers.

Regarding the age of the participants, the topic of filler words is primarily prevalent when taken in regards to teenage speakers, whose vocabularies have evolved to incorporate filler words. I have noticed that in my high school, students use filler words frequently and teachers often include sections in rubrics that dock points for filler words. Also, a common trope in Western media is the portrayal of teenagers overusing common filler words, which can be seen through characters like Shaggy in the *Scooby-Doo* franchise, Leni Loud from *The Loud House*, Mermista in *She-Ra and the Princesses of Power*, Haley from *Modern Family*, and more, which shows that the world has created an association between filler words and teenagers (Like is, Like, a Comma). Since teenagers have been targeted both in the classroom and the media, it is important to not only conduct studies with university students, but to conduct studies with teenage participants to close that gap of knowledge.

Moreover, regarding the participants' linguistic backgrounds, there is a large difference between the speech of a native English speaker and the speech of a non-native English speaker. For example, Odhuu in 2014 cites different "cultural developments" between both parties' patterns of speech. For this reason, it is important to fill the linguistic gap and build a body of knowledge on native English speakers.

Now that the context behind filler words, confidence, and the gaps this study is aiming to fill have been established, it is important to delve into the methods used in this study to answer the question: How does level of confidence affect the frequency in which teenagers apply filler words in their speech?

Methods

This study intended to answer the research inquiry: How does level of confidence affect the frequency in which teenagers apply filler words in their speech? It was designed to be a quantitative, within-subjects design experimental study, assessing the researcher's theory that level of confidence does not affect teenagers' spoken filler word frequency, through numerical data and statistical analyses.

This study is quantitative because the research inquiry does not assess the quality of responses, but trends in the quantities of filler words that make up their speech patterns. Moreover, this study is within-subjects

due to the fact that individuals speak uniquely, so it is only appropriate to compare a participant's data within itself.

Maze Generation

A major part of this study was modeled off of Christenfeld's 1994 study, in which Christenfeld used three types of mazes to inspire conversation for the purpose of recording filler words in their speech. The main difference between his study and this present study is that Christenfeld used three mazes of constant difficulty, while this study's two mazes were changed to produce differing levels of confidence in participants (Christenfeld, 1994). Like Christenfeld, this present study utilized multiple mazes to spark discussion in the participants: Maze A and Maze B (Christenfeld, 1994). As can be seen in Figure 1, Maze A was the simpler maze, designed to produce less self-efficacy feelings and more confidence in the participants, and as can be seen in Figure 2, Maze B was designed to do the opposite.

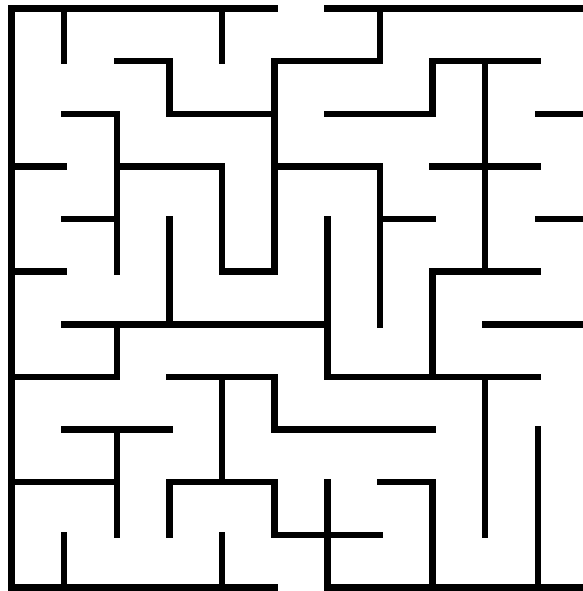


Figure 1. Maze A. As can be seen in Figure 1, Maze A, was generated to produce more confidence with less passageways and obstacles.

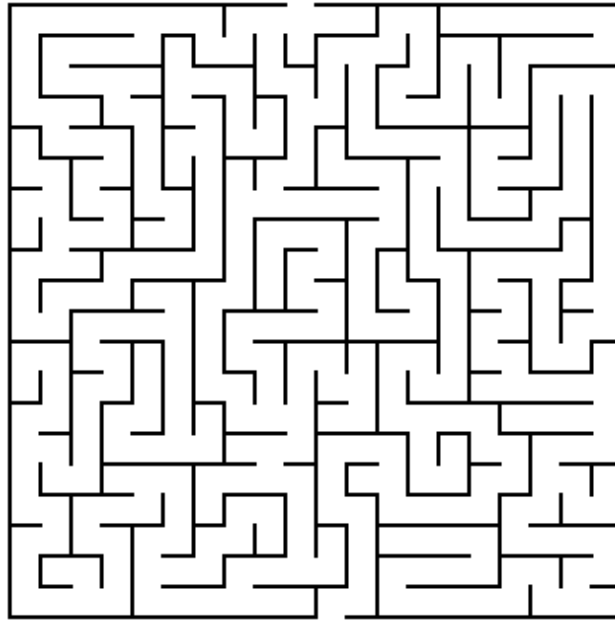


Figure 2. Maze B. As can be seen in Figure 2, Maze B was generated to produce less confidence with more passageways and obstacles.

Ensuring Ethicality

Prior to initiating the first trials of the study, I curated a proposal and consent/ assent forms for this experiment to be approved by the [REDACTED] High School's Institutional Review Board (IRB). The proposal included a description of all predicted aspects of the study, and the consent/assent forms generated for participants and their parents/ guardians to ensure that the study posed no harm to participants or the researcher, and all actions by participants were informed and voluntary (Forms & Templates). The IRB approved of this study, solidifying its adherence to ethical research guidelines.

Participant Recruitment

Convenience sampling was used to gather a sample population for this study. The first step for recruitment was an advertisement stand outside of the [REDACTED] High School's cafeteria during unit lunch time in order to broaden public exposure. Via this stand, students were given an elevator pitch for the study and signed up for an email list to receive study updates.

Before initiating the experiment, the students had to provide written assent and obtain parental consent. These forms were distributed to every student and parent in [REDACTED] High School via the school-sanctioned messaging service. Through this site, students and parents logged on and submitted their forms. After providing both forms, participants waited for the mandatory two weeks before participating in the study, as outlined by the IRB.

For scheduling purposes, I created an interactive calendar using a site called Calendly, which presented available time slots for participants that they could sign up for with a reminder email. On their designated days, students were sent an additional reminder email, providing the information and a reminder regarding their meeting.

Participant Participation

Assignment of a Number Identifier

Primarily, once participants arrived to complete the study, they were assigned with a number which represented their data. Participants were numbered 1-31, in chronological order as they arrived to complete the study. This was done for organization and anonymity.

Completion of a Demographics Questionnaire

Then, participants submitted a demographics questionnaire before moving on to the interactive portion of the study.

I curated this demographics questionnaire to collect data on participants’ basic information, cultural influences, and academic background, all variables that had the potential to skew the results of this experiment.

Viewing and Completion of the Mazes

After collecting their demographic information, participants were presented with two mazes: Maze A and Maze B.

Before solving either maze, participants used Likert Scales to rank their confidence level in completing each maze. To develop this Likert Scale system, aspects of Christenfeld’s study were used in conjunction with Bortfield et al.’s study, who developed a scale for measuring confidence in their participants (Christenfeld, 1994 & Bortfield et al., 2001). Similarly, as can be seen in Table 1, I developed a 5-point Likert Scale which asked participants, after viewing the mazes for the first time, to report their feelings of self-efficacy (Kalra & Siribud, 2020). In reviewing the literature, Likert Scales are questionnaires that are meant to take a self-reported measure of the researcher’s choosing; they are characteristically composed of a scale of one through five, with each number being assigned to an assessment of a statement (Davis & Dymont, 2014).

Table 1. Confidence Likert Scales. Participants were asked to rank their confidence on a scale from 1-5; they were instructed that a score of 1 meant no confidence and a score of 5 meant the most confidence.

I am confident in solving Maze A.				
1	2	3	4	5
I am confident in solving Maze B.				
1	2	3	4	5

For each maze, the participants followed the same procedure. The first maze participants were instructed to solve was Maze A, followed by the discussion session, then Maze B, followed by the discussion session. Although the mazes were of different levels of difficulty, participants were able to complete and explain each in under five minutes.

Participants were allowed to use any method of solving the maze that they preferred, as this study does not focus on the maze completion process itself.

Discussion of the Mazes

Once both mazes were completed, participants moved onto the discussion portion of the study. Since this experiment is a quantitative study, there will be no assessment on the quality of participant responses, but the quantity of filler words within them, so the prompts were framed to inspire conversation and prolong speaking time. Each question has a phrase to spark explanation, such as “in full sentences”, “please elaborate”, and

“why?”; this was to ensure that there was enough material to analyze and the discussion simulated a classroom presentation experience.

Post-Maze Discussion Prompts

1. In full sentences, please describe how you felt when you first saw this maze.
2. In full sentences, please tell me how you solved this maze + what strategies would you suggest for someone who is solving a maze like this?
3. What were the hardest parts of this maze? + what were the easiest parts of this maze? Please elaborate.
4. What age group do you think can solve this maze easily, and why?

Both times, the participants explained their process, prompted with discussion questions, and they were recorded via the Voice Recording computer application. After their recordings were completed, the participant was informed of the true motivation for the experiment (investigating filler words).

Data Analysis

After this observational data was collected, it was analyzed for the amount and type of filler words present. I annotated the participant recording transcriptions for use of filler words, noting both the quantity of filler words, as well as the type of filler word used in each instance. This allows the viewing of participants’ variance or lack of variance in filler word frequency while dealing with Maze A versus Maze B. Identifying the type of filler word helped to determine the cause of the filler word usage. In this study, only the filler words “I mean”, “Like”, “Um”, and “Uh” were identified through the transcriptions, due to model studies like Laserna, 2014 and Bortfield et al., 2001 using the same words, and their popularity.

Addressing the distinction between the two types of filler words is a subsection of the experiment, which tries to assess whether the intention of a filler word used affects the frequency while talking about subjects of varying levels of confidence. As seen in Table 2, every instance of an individual filler word was a color coded highlight on the transcription for easy analysis.

Table 2. Legend for the Color-Coding of Transcriptions.

Filled Pauses		Discourse Markers	
“Um”	“Uh”	“Like”	“I mean”
Red	Orange	Yellow	Green

And finally, the end result was achieved through statistical analysis, where I used Spearman Rank Correlation Coefficient tests and paired *t*-tests to analyze the data fully.

As can be seen by Table 3, each analysis was done at an individual level per participant. I first explored Maze A by finding the percentage of filler words per total word count for Participant 1, and so on until Participant 31. After this, I found the total sum of the percentages for each participant. Then, to find the average percentage of filler words to total word count for Maze A, I divided the total sum of percentages by the total number of participants in the study, which was 31 participants.

Table 3. Percentage of Speech Consisting of Filler Words for Each Maze.

Maze A			Maze B		
Sum of Percentages of Filler Words to Total Words	Total Participants	Average Percentage of Filler Words to Total Words	Sum of Percentages of Filler Words to Total Words	Total Participants	Average Percentage of Filler Words to Total Words
171.805	31	5.542%	176.324	31	5.688%

Then, while annotating participant transcriptions, I also observed the quantity of filled pauses and discourse markers used in order to identify any great differences in these variables between confidence events. To find this for each maze, I individually took the percentage of filled pauses to total filler words for each participant and then did the same for discourse markers to total filler words for each participant. Then, I divided the sum of each collection of percentages and by the number of participants (31) to find the average percentage of filled pauses to total filler words for each maze and average percentage of discourse markers to total filler words for each maze.

Once the quantitative analysis was reached, I was able to determine the effect of confidence on filler word usage in teenagers using the following logic: if the participants consistently employed more or less filler words when speaking about one maze, then it could be concluded that filler words have a correlation with confidence. However, if the participants consistently employed the same amount of filler words for each maze, then it could be assumed that confidence did not have an effect on filler word usage.

Results

Sample Population

To address the research question of this study, “How does teenager’s confidence affect their use of filler words in their speech?”, this experiment was conducted on 31 students from █████ High School, a single public high school, and as can be seen by Figure 3 and Figure 4, the students were aged 14-18 years old and of varied genders.

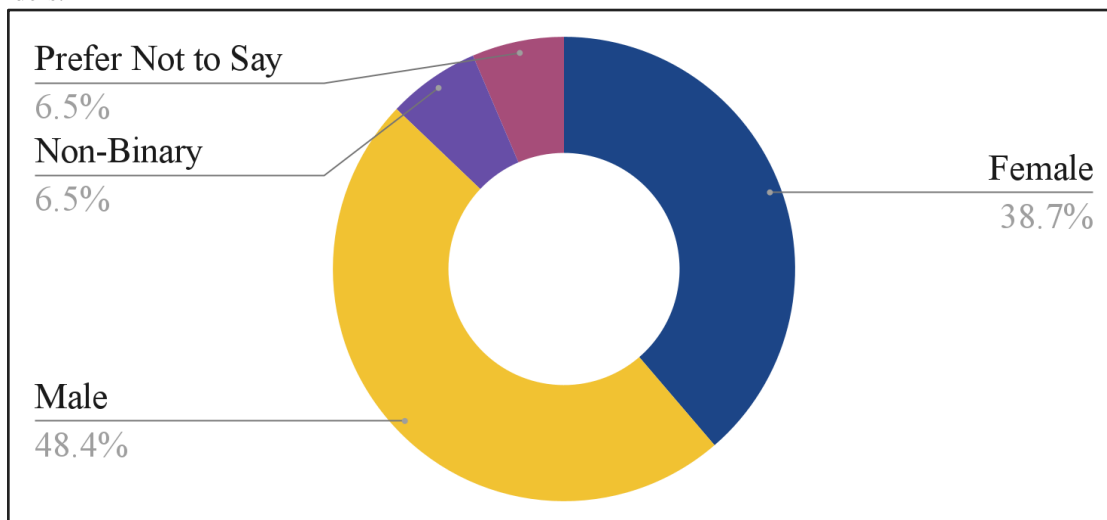


Figure 3. Participant Gender Distribution.

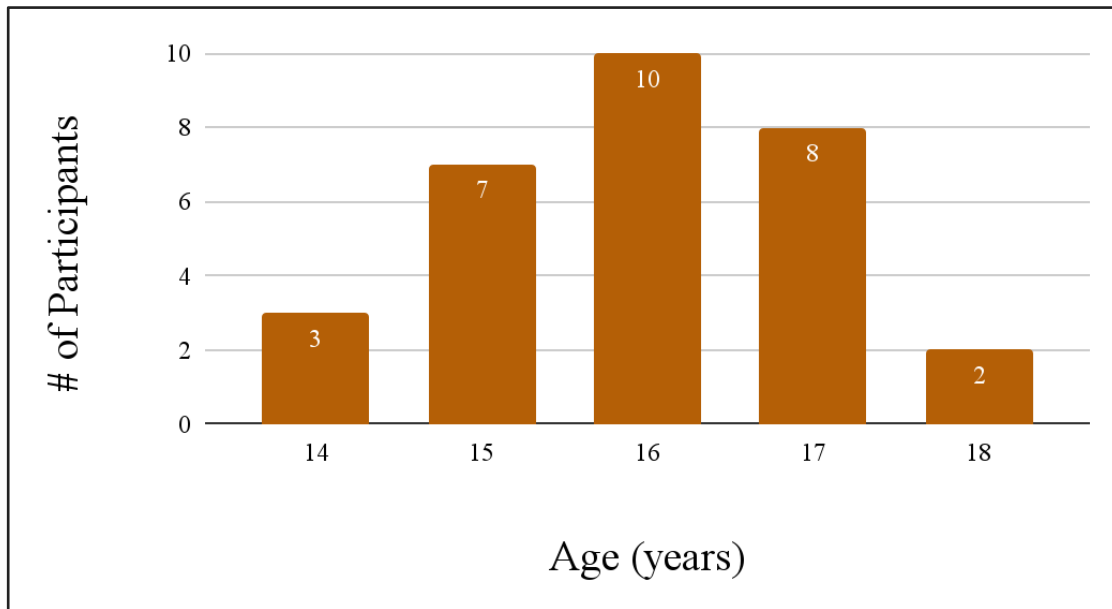


Figure 4. Participant Age Distribution

Likert Scale Results

The results of the Likert Scale submissions by participants show that the average level of confidence for Maze A was 4.7 and the average level of confidence for Maze B was 3.8.

The data sets used in this test are from the same population before and after the mazes changed and diminished their confidence. For this reason, as outlined by David Borman, paired/ dependent samples *t*-testing was the most appropriate test to use to test the difference in means of the two variables (Borman, 2018).

As can be seen in Table 4, results of the paired *t*-test performed on the two data sets indicated that there is evidence of a significant difference between the mean confidence levels submitted for Maze A ($M = 4.7$, $SD = 0.5$) and Maze B ($M = 3.8$, $SD = 0.9$), $t(30) = 7.7$, $p < .001$. Therefore, it is appropriate to assert that Maze A and Maze B are mazes that evoke different levels of confidence from participants.

Table 4. Paired samples *t*-test data for participant confidence in Maze A and participant confidence in Maze B.

Variable	Mean	Standard deviation	Standard error of the mean	<i>t</i> (30)	<i>p</i>
Maze A	4.709	0.4614	0.0828	7.6599	< 0.001
Maze B	3.774	0.9204	0.1653		

The Strength of Relation Between Confidence and Filler Word Frequency

To answer the research question, “How does level of confidence affect the frequency in which teenagers apply filler words in their speech?”, it was imperative to investigate the strength of the relation between the two

prominent variables in the study: level of confidence (independent variable) and percentage of filler words employed (dependent variable). The level of confidence in a participant was determined using a 5-point Likert Scale, which means that the data is discrete and ordinal, not normally distributed (Urdan, 2005). On the other hand, the percentage of filler words employed within a participant’s speech, does not have definitive integers to adhere to, so the data is continuous and normally distributed (Urdan, 2005). In addition, the variables used in this test are paired; each participant’s confidence level value corresponds to that same participant’s percentage of filler words. As defined by Dr. Harry Khamis, due to one variable being ordinal and the other being continuous, a non-parametric correlation coefficient test was needed to assess the strength of their relation; so, it was most appropriate to perform a Spearman Rank Correlation Coefficient test—a modified form of the parametric Pearson Correlation Coefficient test—to assess the strength of relation between the data sets (Khamis, 2008).

Following the test procedure outlined by Mavuto Mukaka, a Spearman Rank Correlation Coefficient test was computed to analyze the relationship between level of confidence and filler word frequency. Based on the scale defined by Makaka, which can be seen in Table 5, on how to interpret the *r* value in relation to strength of correlation between two variables, there is a negligible correlation between the two variables, $r(29) = -.1, p = .184$ (Makaka, 2012).

Table 5. Mavuto Makaka’s 2012 scale for interpreting correlation coefficient tests.

Size of Correlation	Interpretation
.90 to 1.00 (-.90 to -1.00)	Very high positive (negative) correlation
.70 to .90 (-.70 to -.90)	High positive (negative) correlation
.50 to .70 (-.50 to -.70)	Moderate positive (negative) correlation
.30 to .50 (-.30 to -.50)	Low positive (negative) correlation
.00 to .30 (.00 to -.30)	Negligible correlation

Within the bounds of this study and its sample population, the result of this correlation coefficient test implies that a varying level of confidence will not correlate to a varying level of filler words spoken in teenage speech, so this supports the stance that a teenager’s level of confidence in a subject will not affect the quantity of filler words in their speech.

The Statistical Significance of the Difference in Means of Participant Filler Word Frequency for Each Maze

In order to compare the means of participants’ percentages of filler words in each Maze to gauge statistical significance, a *t*-test was needed. The data sets used in this test are from the same population before and after their mazes changed and diminished their confidence. And additionally, each filler word frequency from Maze A corresponded with the same participant’s filler word frequency from Maze B, so the data is paired. For this reason, paired/ dependent samples *t*-testing was the most appropriate test to use to test these variables (Borman, 2018). The null hypothesis established for this test was if confidence is decreased in teenagers, then their filler word usage will remain the same. Conversely, the alternative hypothesis established for this test was if confidence is decreased in teenagers, then their filler word usage will change. As can be seen in Table 6, using an accepted significance level of 0.05, the results of the paired *t*-test indicated that there is no convincing evidence of a difference in the means of the filler word frequency data before level of confidence is decreased ($M = 5.5, SD = 2.8$) and the filler word frequency after level of confidence is decreased ($M = 5.7, SD = 3.1$), $t(30) = .3, p = .740$. Thus, the null hypothesis, the assumption that if confidence is decreased in teenagers, then their filler word usage will remain the same, cannot be rejected within the bounds of this study.

Table 6. Paired samples t-test data for level of confidence and filler word frequency.

Variable	Mean	Standard deviation	Standard error of the mean	t(30)	p
Maze A	5.542	2.7770	0.4988	0.3350	0.7399
Maze B	5.688	3.0784	0.5529		

Delving into The Strength of Relation Between Confidence and Frequencies of Different Filler Words

In addition, to delve into a subsection of the original research question, which investigates the relationship between specific types of intended filler words (filled pauses and discourse markers) and level of confidence, it was most appropriate to perform a Spearman Rank Correlation Coefficient test for each variables' relations to level of confidence. This is because both of these tests involved the same ordinal, non-normally distributed Likert Scale data that was used in the first Spearman Rank Correlation Coefficient test; any data set that involves a non-continuous variable cannot be tested using a parametric test, such as the commonly used Pearson Correlation Coefficient Test (Khamis, 2008).

The Strength of Relation Between Confidence and Filled Pause Frequency

The strength of the relation between level of confidence (independent variable) and discourse marker frequency (dependent variable) was then tested. The variables used in this test are paired; each participant's confidence level value corresponds to that same participant's percentage of discourse markers spoken. A Spearman Rank Correlation Coefficient test was computed to analyze the relationship between level of confidence and filled pause frequency. Based on the scale defined by Makaka, which can be seen in Table 5, there was a negligible correlation between the two variables, $r(29) = .2$, $p = .029$ (Makaka, 2012). This implies that within the bounds of this experiment, a varying level of confidence will not correlate to a varying level of filled pauses spoken in teenage speech, so the intentions behind filled pauses are not a factor when considering the research inquiry.

The Statistical Significance of the Difference in Means of Filled Pause Frequency for Each Maze

In order to compare the means of participants' percentages of filled pauses in each Maze to gauge statistical significance, a *t*-test is needed. The data sets used in this test are from the same population before and after their confidence diminished. And additionally, each filled pause frequency from Maze A corresponded with the same participant's filled pause frequency from Maze B, so the data is paired. For this reason, paired/dependent samples *t*-testing is the most appropriate test to use to test these variables (Borman, 2018). The null hypothesis established for this test is if confidence is decreased in teenagers, then their filled pause usage will remain the same. Contrastingly, the alternative hypothesis established for this test is if confidence is decreased in teenagers, then their filled pause usage will change. As can be seen in Table 7, using an accepted significance level of 0.05, the results of the paired *t*-test indicate that there is no convincing evidence of a difference in the means of the filled pause frequency data before level of confidence is decreased ($M = 55.1$, $SD = 27.7$) and the filled pause frequency after level of confidence is decreased ($M = 50.8$, $SD = 30.1$), $t(30) = .8$, $p = .412$. Thus, the

null hypothesis, the assumption that if confidence is decreased in teenagers, then their filled pause usage will remain the same., cannot be rejected within the bounds of this study. So, this result solidifies that within the bounds of this study, the intentions behind filled pauses are not a factor when considering the research inquiry.

Table 7. Paired samples t-test data for filled pause frequency before and after level of confidence decreased.

Variable	Mean	Standard deviation	Standard error of the mean	t(30)	p
Maze A	55.091	27.6996	4.9750	0.8325	0.4117
Maze B	50.824	30.1313	5.4117		

The Strength of Relation Between Confidence and Discourse Marker Frequency

Then after testing the values of confidence and filled pause usage, the strength of the relation between level of confidence (independent variable) and discourse marker frequency (dependent variable) was tested; data classifications are unchanged from the test for filled pauses. A Spearman Rank Correlation Coefficient test was computed to analyze the relationship between level of confidence and discourse marker frequency. Based on the scale defined by Makaka, which can be seen in Table 5, there was a negligible correlation between the two variables, $r(29) = -.2, p = .110$ (Makaka, 2012). This implies that within the bounds of this experiment, a varying level of confidence will not correlate to a varying level of discourse markers spoken in teenage speech, so the intentions behind discourse markers are not a factor when considering the research inquiry.

The Statistical Significance of the Difference in Means of Discourse Marker Frequency for Each Maze

In order to compare the means of participants' percentages of discourse markers in each Maze to gauge statistical significance, a *t*-test is needed. The data sets used in this test are from the same population before and after their confidence lessened. And additionally, each discourse marker frequency from Maze A corresponded with the same participant's discourse marker frequency from Maze B, so the data is paired. For this reason, paired/dependent samples *t*-testing is the most appropriate test to use to test these variables (Borman, 2018). The alternative hypothesis established for this test is if confidence is decreased in teenagers, then their discourse marker usage will change. Contrastingly, the null hypothesis established for this test is if confidence is decreased in teenagers, then their discourse marker usage will remain the same. As can be seen in Table 8, using an accepted significance level of 0.05, the results of the paired t-test indicate that there is no convincing evidence of a difference in the means of the discourse marker frequency data before level of confidence is decreased ($M = 44.9, SD = 27.7$) and the discourse marker frequency after level of confidence is decreased ($M = 45.9, SD = 29.9$), $t(30) = .3, p = .797$. Thus, the null hypothesis, the assumption that if confidence is decreased in teenagers, then their discourse marker usage will remain the same, cannot be rejected within the bounds of this study.

Table 8. Paired samples t-test data for discourse marker frequency before and after level of confidence decreased.

Variable	Mean	Standard deviation	Standard error of the mean	t(30)	p

Maze A	44.909	27.6996	4.9750	0.2596	0.7970
Maze B	45.950	29.8605	5.3631		

Strengths of Relations & Statistical Significances of Filled Pauses and Discourse Markers Means

In all, as can be seen by the inferential statistics on both types of filler words, within the bounds of this study, it can be concluded that the type of filler word used does not have a relation to the level of confidence in a subject. Therefore, while a filled pause may signal a different intention from the speaker than does a discourse marker, this distinction is not affected by the level of confidence in the teenager speaker within the bounds of this experiment.

Discussion

This study set out to answer the question: How does level of confidence affect the frequency in which teenagers apply filler words in their speech? This was performed via a quantitative experiment with a sample population of 31 high school students. Ultimately, as outlined by the researcher’s original hypothesis, which is that a teenager’s level of confidence in a topic does not affect their frequency of spoken filler words, a negligible relationship was found between the level of confidence in a teenager and their spoken filler words, within the bounds of this study. Additionally, the type of filler word used—filled pauses versus discourse markers—was not found to have had a significant effect on this relationship.

However, before applying the conclusions of this study to the general population, there are several limitations to this research that are important to address.

This study is limited in age representation due to the limitations of convenience sampling, as can be seen in Figure 4, which shows that 32% of the sample population were 16 year olds, 9% of participants were 14 years old, and 6% were 18 years old. These imbalances may skew data, since participants of different ages have been introduced to more styles of speaking and have completed more language arts classes that may have conditioned them to eliminate filler words.

In addition, studies have shown that men are more likely to overuse filler words in their speech. Bortfield et al. discovered that males, within the scope of their experiment, produced 2.57 more filler words per every 100 words than their female counterparts (Bortfield et al., 2001). Considering that 48.4% of the sample population was comprised of male participants, as can be seen in Figure 3, this factor could have skewed results, as male participants would be more disposed to use filler words than female or nonbinary participants, regardless of their confidence level.

Also, non-native English speakers have different cultural developments linguistically (Odhuu, 2014). For this reason, it was a goal to have a native English speaker population, however, due to the limitations of convenience sampling, as can be seen in Figure 5, 22.6% of participants were non-native English speakers, which may have introduced skewed data.

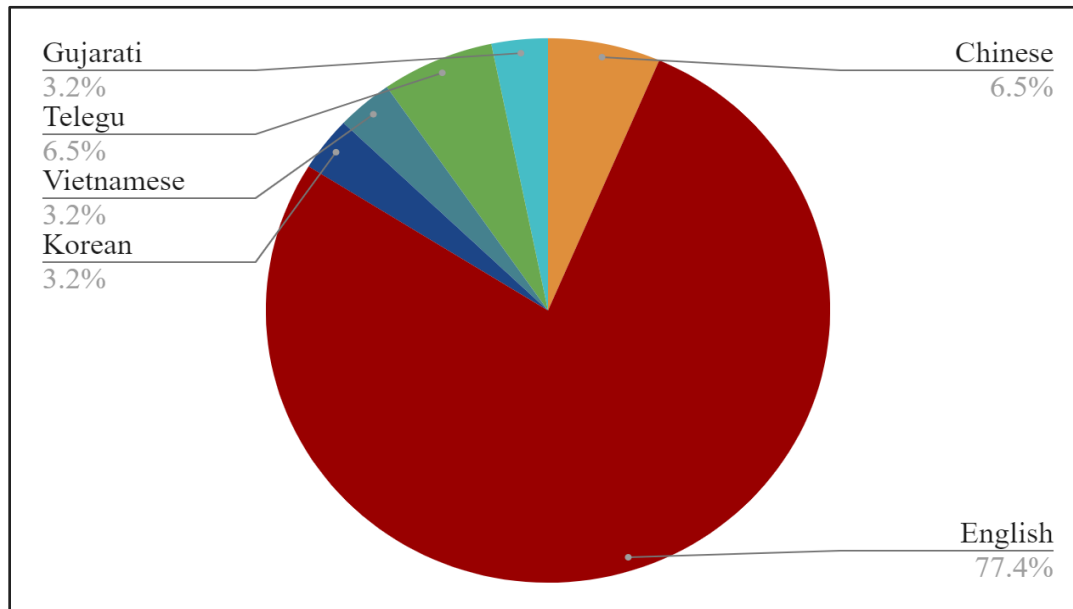


Figure 5. Participant First Language Distribution.

Further research should identify whether other factors in the study can create a disruption to filler word frequency and confidence level. For example, studying whether the location of the filler word within a sentence is impactful when analyzing patterns. In addition, the action of self-ranking confidence within participants should be explored and refined, to fine-tune the measure of confidence in a situation. The setting of a situation can also be explored in order to control that variable of the experiment, such as the volume of the environment, or the number of people present while the participants are speaking. And overall, population limitations that came about due to convenience sampling, such as age, gender, and language, should be addressed.

The negligible correlation between confidence and filler word frequency found within this study could have further implications in the world of high school education and grading systems in classrooms. If developed further, the data presented could contribute to breaking down the negative connotation surrounding filler words and their relation with confidence. For example, the reform of harsh grading systems and curriculums in place for classrooms and representation for teenagers in the media.

Acknowledgments

I would like to thank my advisor for the valuable insight provided to me on this topic.

References

- Alvar, A. M., Huber, J. E., & Lee, J. (2019). Filled pauses as a special case of automatic speech behaviors and the effect of parkinson's disease. *American Journal of Speech - Language Pathology*, 28(2S), 835-843. http://dx.doi.org/10.1044/2018_AJSLP-MS18-18-0111
- Battistella, E. L. (2007, August 22). There's a lot to like. Retrieved December 6, 2021, from <https://web.s.ebscohost.com/ehost/detail/detail?vid=9&sid=f4f32ccc-a923-4d4e-838b-30eb36e8771c%40redis&bdata=JnNpdGU9ZWZWhvc3QtbGl2ZQ%3d%3d#AN=26322930&db=lfh>
- Borman, D. (2018). *Statistics 101: from data analysis and predictive modeling to measuring distribution and determining probability, your essential guide to statistics* (2018 ed.). Adams Media.

- Bortfield, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender. *Language and Speech, 44*(2), 123-147. <https://doi.org/10.1177%2F00238309010440020101>
- Chang, C.-C. A., & Yang, X. (2010). Confidence, mindset, and self-efficacy in goal pursuit. *Advances in Consumer Research, 37*, 767-768.
- Christenfeld, N. (1994). Options and ums. *Journal of Language and Social Psychology, 13*(2), 192-199. <https://doi.org/10.1177%2F0261927X94132005>
- Crible, L., & Kosmala, L. (2021). The dual status of filled pauses: evidence from genre, proficiency and co-occurrence. *Language and Speech, 1*-24. <https://doi.org/10.1177%2F00238309211010862>
- Davis, J. M., Dymont, J. E., Emery, S., Getenet, S., Hill, A., & McCrea, N. (2014). The impact of professional development on early childhood educators' confidence, understanding and knowledge of education for sustainability. *Environmental Education Research, 660*-679. <http://dx.doi.org/10.1080/13504622.2013.833591>
- Forms & Templates*. (n.d.). Smith College. Retrieved April 29, 2022, from <https://www.smith.edu/academics/institutional-review-board/forms>
- Fox Tree, J. E., & Schrock, J. C. (2002). Basic meanings of you know and I mean. *Journal of Pragmatics, 34*(6), 727-747. [https://doi.org/10.1016/S0378-2166\(02\)00027-9](https://doi.org/10.1016/S0378-2166(02)00027-9)
- Gósy, M., & Horváth, V. (2010). Changes in articulation accompanying functional changes in word usage. *Journal of the International Phonetic Association, 40*(2), 135-161. <http://dx.doi.org/10.1017/S0025100310000058>
- Kalra, R., & Siribud, S. (2020). Public speaking anxiety in the Thai EFL context. *Language Education and Acquisition Research, 13*(1), 195-209.
- Khamis, H. (2008, March 14). Measures of association: how to choose? *SAGE Journals Online*. Retrieved April 28, 2022, from <https://doi.org/10.1177%2F8756479308317006>
- Laserna, C. M., Pennebaker, J. W., & Seih, Y.-T. (2014). Um . . . who like says you know: filler word use as a function of age, gender, and personality. *Journal of Language and Social Psychology, 33*(3), 328-338. <https://doi.org/10.1177%2F0261927X14526993>
- Like is, like, a comma*. (n.d.). TV Tropes. Retrieved May 1, 2022, from <https://tvtropes.org/pmwiki/pmwiki.php/Main/LikeIsLikeAComma>
- Like like like like what?* (2015). SiOWfa15: Science in Our World: Certainty and Controversy. Retrieved April 30, 2022, from <https://sites.psu.edu/siowfa15/2015/12/02/like-like-like-like-what/>
- Makaka, M. (2012, September). *A guide to appropriate use of correlation coefficient in medical research*. National Library of Medicine. Retrieved April 28, 2022, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3576830/>
- Odhoo, K. (2014). Teaching the native english speaker how to teach english. *English Teaching Forum, 52*(2), 2-11.
- Urdu, T. C. (2005). *Statistics in plain english* (2nd ed.). Psychology Press.