

2014-01-01

Language Difference Between Peer And Professional Drug And Substance Abuse Counselors: What Can It Tell Us?

Lizzeth Celeste Jensen

University of Texas at El Paso, lizspeedvelmag@gmail.com

Follow this and additional works at: https://digitalcommons.utep.edu/open_etd



Part of the [Anthropological Linguistics and Sociolinguistics Commons](#)

Recommended Citation

Jensen, Lizzeth Celeste, "Language Difference Between Peer And Professional Drug And Substance Abuse Counselors: What Can It Tell Us?" (2014). *Open Access Theses & Dissertations*. 1266.

https://digitalcommons.utep.edu/open_etd/1266

LANGUAGE DIFFERENCE BETWEEN PEER AND PROFESSIONAL DRUG
AND SUBSTANCE ABUSE COUNSELORS:
WHAT CAN IT TELL US?

LIZZETH C. JENSEN

Department of Languages and Linguistics

APPROVED:

Jon Amastae, Ph.D., Chair

Nicholas Sobin, Ph.D.

Mary Cuadrado, Ph.D.

Charles Ambler, Ph.D.
Dean of the Graduate School

Copyright ©

by

Lizzeth C. Jensen

2014

Dedication

To my daughter Natasha, may you never lose the joy of learning.

LANGUAGE DIFFERENCE BETWEEN PEER AND PROFESSIONAL DRUG
AND SUBSTANCE ABUSE COUNSELORS:
WHAT CAN IT TELL US?

by

LIZZETH C. JENSEN, B.A.

THESIS

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF ARTS

Department of Languages and Linguistics

THE UNIVERSITY OF TEXAS AT EL PASO

December 2014

Acknowledgements

I would like to acknowledge and give thanks to the following people who played a huge role in the creation and completion of this thesis:

I would like to begin with my mentors, Dr. Jon Amastae and Dr. Mary Cuadrado who were there for me through this long and tedious process every step of the way.

Dr. Amastae, thank you for always sharing your knowledge and passion of linguistics and the hocus pocus therein and for thinking of me for the great opportunity of working under the VIDA project.

Dr. Cuadrado, thank you for showing me the ropes and getting my feet wet in the Drug Addiction field. For sharing your knowledge with me, for not letting me give up, and for giving me tough love when you knew I needed it. Above all thank you for being there for me through everything that happened in the years we worked together.

To the Vulnerability Issues in Drug Abuse (VIDA) program family: From Dr. Castañeda to Rosemary and all the RA's just like me, thank you for doing the best you could to help me grow as a researcher and as an individual.

To Dr. Sobin who took the time to be a part of this committee and for always sharing his passion of Syntax and Computational linguistics.

To Dr. Maria Blume for always being such a positive motivator and always telling me to pursue and continue with my education.

To my loving husband John: thank you for the encouraging words you always knew to say when I was ready to pull my hair out. Not letting the distance of you being deployed stand in the way of your support for me in this endeavor.

Para mi mai: gracias por el apoyo que me has brindado durante esta odisea, especialmente con el cuido de Natty.

To my princess Natasha, thank you for doing your best in giving mommy time to be able to write and for giving up sooo many hours of playtime with me so I could get this done. I love you.

To my wonderful friends in the LALR Lab who always cheered my on. Especially Martha, Thank you so much for helping me without hesitation with my reliability checking.

To the Ling Crew (Lilli, Hasan, Alexis, Marty, Monica, Laura, Raquel, Yumira and Stephen) I miss you guys and thank you for being examples for me to follow with your own Theses, extended papers and presentations. For always putting a smile on my face when I was down and discouraged, for helping me practice and edit. But above all for being the amazing Friends you are.

Without you all none of this would have been possible. Thank you.

Table of Contents

Acknowledgements.....	v
Table of Contents	vii
List of Tables	x
List of Graphs	xi
Chapter 1: Language and State of Mind	1
1.0 Introduction.....	1
1.1 Origin	1
Chapter 2: Previous Studies	3
2.0 Introduction.....	3
2.1 Peer v Professional Counselor Studies	3
2.2.0 Language and Behavior of Mind	5
2.2.1 Language and Behavior Testing Methods	6
2.3 Language and Treatment.....	9
2.4 Limitations of Previous Studies	12
Chapter 3: The Present Study	13
3.0 Introduction.....	13

3.1 Research Questions	13
3.2 Sample and Data Collection.....	14
3.3 Definitions and Distinctions	15
3.4 Process	15
3.4.1 Reliability Checking	15
3.4.1.1 Issues with Recordings/Transcripts	16
3.4.2 Language Analysis Programs	17
3.4.2.1 AntConc	17
3.4.2.2 Linguistic Inquiry and Word Count.....	19
3.5 Order	20
3.6 Merging Changes in Transcripts.....	20
3.7 Preparing Transcripts for Analysis	21
3.8 Sample.....	22
3.8.1 Description of the Sample.....	22
3.8.2 Sample Excerpts from Interviews.....	23
3.8.2.1 Peer Sample	23
3.8.2.2 Pro Sample	23
Chapter 4: Results	24
4.0 introduction	24

4.1 AntConc	24
4.1.1 Results	24
4.1.2 Discussion of AntConc Findings	25
4.2 LIWC	26
4.2.1 Results	26
4.2.2 SPSS	27
4.2.2.1 Pronouns	29
4.2.2.2 Psychological Processes	36
4.2.2.3 Other Linguistic Categories	39
4.2.2.4 All Other Categories	42
4.3 Integration	43
Chapter 5: Discussion and Conclusion	45
5.0 Introduction	45
5.1 Previous Research and Current Findings	46
5.2 Conclusion	47
5.3 Limitations of Study	49
Bibliography	50
Appendix	52
Vita	62

List of Tables

Table 7 T-test significance of data.....	29
Table 8 Pro Pronoun use and mean.....	30
Table 9 Peer Pronoun use and mean.....	31
Table 10 Different types of “you” Peer Counselors.....	34
Table 11 Different types of “you” Pro Counselors.....	35
Table 12 Pro Linguistic Categories.....	39
Table 13 Peer Linguistic Categories.....	40

List of Graphs

Graph 1 Pronoun use differences between Peer and Pros.....	32
Graph 2 Psychological Process word use difference between Peer and Pros.....	37
Graph 3 Psychological Process Peer v Pro continued.....	38

Chapter 1: Language and State of Mind

1.0 Introduction

“[H]uman thinking is governed on the level of states of mind, but not on the level of the mind itself. In other words, the thinking is governed indirectly that influences essentially its laws, e.g., the character of the logic used by human beings [...]” (YaTemkin, 2011). YaTemkin, made the connection of how human thinking is on a different plane than the regular mind. By saying there are “states of mind” it shows that when we try to know what someone’s thinking, we have to find the correct state that we are looking for, but how? Consequently, language has been thought to be a tool to reveal the state of a person’s mind. The purpose of this study was to:

- i) initiate study involving the analysis of difference between peer and professional counselors
- ii) investigate how the language of peer counselors in comparison to professional counselors could reveal a difference in the state of mind of each
- iii) analyze the efficacy of language analysis tools to analyze natural speech without human intervention. Through the use of two different language analysis tools the answers to these questions were sought.

1.1 Origin

In recent years, new approaches to empirical investigation of language and mind have proliferated. None have yet examined the topic of this thesis, which is whether, through the use of language, it is possible to detect a difference between two groups of people engaged in the same professional activity, in this case peer counselors compared to professional counselors and their language on substance addiction treatment.

This research question arose while performing interviews for the project “The Use of “Juramentos” As a Substance Abuse Social Control Mechanism” under the Vulnerability Issues in Drug Abuse (VIDA) program led by Dr. Mary Cuadrado. In the Juramento project, peer-led directors would tell the interviewers “no offense, but I didn’t get my experience from books.” While others made sure to clarify that their counselors were professionally trained. Inquiries into whether this was a common phenomenon revealed an informal rivalry, where peer counselors perceive themselves to be better suited to provide treatment due to their personal experience and their personal success in overcoming addiction.

This study brings together three areas: 1) Peer v. Professional counselors, 2) studies of language and behavior, and 3) studies that deal with language use in clinical contexts. Presently, there exist no studies that incorporate all three areas in one. This study hopes to connect the study of language of Peer and Professional counselors of Substance use disorder to see if a difference exists amongst the two that could account for why this notion of Peer counselors being better at what they do, exists.

Chapter 2: Previous Studies

2.0 Introduction

The study of language and what it reveals of a person's state of mind is not a new phenomenon. However, no study has sought to tackle the apparent rivalry between Peer and Professional counselors, as to whether experience versus education translate into a different approach with dealing or expressing oneself about addiction. New techniques of analysis give way to be able to both explore a practical professional issue (i.e. Peer versus Pro) and be able to contribute something to an old general question of: Does language reveal the mind? And if so, how?

2.1 Peer v Professional Counselor Studies

In the realm of drug/substance abuse treatment there is an apparent rivalry that either peer-led or professional-led treatment approaches are better than the other. Studies looking at effectiveness of peer or ex-addicts as counselors have been in existence well before the 1960s, and continue to be debated to this day (Kadushin & Kadushin, 1969). Both sides believe that what each type of counselor brings to the treatment table or what they offer the patient is the experience they learned; either from educational experience, or personal experience, is better. Up to now, there has been no empirical investigation of the many possible differences in thinking or approach of peer versus professional counselors, particularly in regard to their language and the possible differences therein.

For the Professional versus Peer studies, articles that speak about Peer Led treatments and professional led treatments were chosen, to have an idea of what each type of counselor is stated

to represent according to literature. What was found was a mixture of positive and negative thoughts of what an ex-addict counselor can provide to treatment and the risks that the peer counselor himself can face by involving themselves in addiction recovery programs. A 1969 article by Kadushin and Kadushin, goes into detail of how an ex-addict should be incorporated into treatment plans and the positive of incorporating them. Yet, it also mentions how ex-addicts should still be supervised to ensure that the integrity of their treatment program could be kept. Kadushin and Kadushin (1969) state “the ex-addict frequently may be the only member of the treatment team with whom the addict can communicate, especially at the onset of treatment. He is often the motivating force propelling the addict to enter a treatment program and, once there, may help the addict stick it out” (p388). Other articles from the same era mention the positive of hiring ex-addicts as group leaders, by being cheaper to hire, and they could serve as a gateway to the addict because the ex-addict had the validation from addicts in term of them knowing what the addict was experiencing. (Kadushin & Kadushin, 1969) (Snowden Jr & Cotler, 1974). However, as clearly stated in the article by Doukas and Cullen (2010) just how the literature of the effectiveness and the positive side of ex-addicts in counseling starting the in the 1940s bloomed, by the 1990s it began to subside and hardly any talk of it continued in the 21st century.

Doukas and Cullen’s article is an extensive literature review of all the literature that has been available since the 1940s up to the year 2010 regarding peer counselors in the field of addiction recovery. The article lists the downside of “recovered counselors” and backs it up with literature. Some of the downsides are “recovered counselors are inflexible in their approach by virtue of being resistant to new learning (Delali, Charuvastra and Schlesinger 1976; Siassi, Angele & Alston 1977), being overcommitted to one treatment modality due [to] a personal loyalty to the 12 step approach (Humphreys, Noke & Moos 1996; Shipko & Stout 1992;

Stoffelmayr et al 1998) and operating from a limited frame of reference because they view all clients in terms of being addicted or not, which might lead to over diagnosis (Lawson, Petosa & Peterson 1982).” (Doukas & Cullen, 2010, p. 217) Doukas and Cullen continue by saying “an obstacle commonly faced in the helping profession is when counselors cannot detach themselves from the clients they work with and become overly involved with their recovery, often taking on the role of a rescuer (Valle 1979).” (p.221). This article was pertinent to this study because it shows from a point of view of a recovered counselor, or as this study will refer to them as peers, what they think about the recovery field and the downsides they may encounter when faced with counseling. It must also be noted that the articles found pertaining to peer and professional counselors did not mention language differences in everyday speech like this study is trying to compare.

2.2.0 Language and Behavior of Mind

Language and behavior have both been known to be the most obvious ways to capture the mind. How can these two different approaches though work together? In order to better understand these forms we must first know what language and behavior have in common. An enigma that many have sought to explain is finding out what is in other people’s minds that may explain their behavior. Language has given us a closer look into people’s minds, but, the language we share with one another does not necessarily say what is truly on our mind. There are many reasons why language can be tricky when it comes to using it as a tool for sharing state of mind; some of the things are: 1) the person will just tell you what they want you to hear, aka Lies; 2) self-deception, where the person believes it's true but it is not; and 3) lack of self-awareness. How can these problems be overcome to get an accurate reading of a person’s state of mind? That is what many researchers have tried to figure out by using different methods.

2.2.1 Language and Behavior Testing Methods

In the extended research paper by Laura Mendoza, a method of testing the way that language's syntactic and semantic structure can show one's mind is introduced. Mendoza's study consisted of using the UTEP's Learner Corpus of Academic English (ULCAE) in which data had been inputted or transcribed in, ending in a total of "1,355,645 words". These words, as Mendoza (2012), states "reflect[s] a variety of writing genres such as process essays, comparison and contrast essays, evaluative essays and research reports, among others." (p.9). In order to analyze the massive corpus, Mendoza used a program called MonoConc Pro to help her find the "essays to be included in the 'v' sub corpus, the automated text search feature of the concordance software 'MonoConcPro' was used to locate the presence of any of the target items (from Urzúa & Mendoza 2011) in the original data set." (Mendoza, 2012). The first time Mendoza ran the data through the program, all that was searched for were essays with sensitive topics. Later on, the data was broken down further into essays that contained topics pertaining to "social conflict and/or violence" (Mendoza, 2012, p. 11). In her paper, Mendoza (2012) looks at "determining patterns of lexical co-occurrence" and "determin[ing] any patterns in the use of linguistic markers of stance expressed in relation to violence and conflict being described by the students."(p.1). In other words, Mendoza saw in her ESOL classes, how students would continuously bring up violence related topics even though they were beginning English speakers and their English lexicon was limited. This was particularly important or, strange because these students were from different economic backgrounds living in Cd. Juárez and now studying in El Paso, TX and when given the choice of writing whatever they wanted for their assignment, they would revert to speaking about the violence that had affected them in some way, shape or form, occurring in Cd. Juárez, Mexico. Regardless of their lack of knowledge of the new language the

students still felt the need to share their experiences as best as they could because it had affected them so much in their daily lives. This goes to prove that humans are social beings and through language, even if they are not fluent in the target language humans still want to share their experiences with others.

Although revealing what is inside of us goes back to the Greeks, modern approaches have started to add computers to the mix of finding out what is inside the black box of the mind. Aside from the study previously described that used MonoConc Pro, Dr. James Pennebaker's studies analyze specific words that are used in regular everyday text, some elicited some not. He does this through the use of the Linguistic Inquiry and Word Count (LIWC for short) through which he takes very specific words or phrases that can be emotionally heavy or gender specific. These words are analyzed to give an interpretation of the character of that person (Tauszcisk, 2010). As best described in his book *The Secret Life of Pronouns: What our words say about us* the production of the program LIWC was delayed three years because of the vast dictionaries created for it that Pennebaker wanted to capture "different psychological concepts" (Pennebaker, 2011, p. 6) as well as the process it took to be able to build each dictionary. Pennebaker states that he "employed an army of students who evaluated every word that was part of any dictionary [and...] the judges had to all agree that it was related to [the dictionary category]." (Pennebaker, 2011). Ultimately though, after the creation of LIWC, the program is supposed to count all the words in the text, and then follow the word count by looking at each individual word to determine if the word belongs to any of the pre-determined dictionaries. If the word that is being analyzed does not fit any of the dictionaries then the program would move on to the next word. Doing so for each word found in the text and trying to pair each word within a dictionary (or pairing it with multiple dictionaries if the word could fit into more than one category/dictionary).

If, the program succeeded in matching one word with one distinct category/ dictionary the word count for that dictionary would go up to 1. After the analysis of the whole text is completed then LIWC calculates the percentage of “total words that are linked to each dictionary” (Pennebaker, 2011, p. 8) The pros of this program is that it can analyze a large sized text in a matter of seconds and can give you an idea of the mentality and overall pattern of use of the speaker or speakers. It also give the advantage that since all the categories stay the same for each file analyzed, it is much easier to compare results amongst speakers or texts or even within different texts of the same speaker but from different time periods. The cons of LIWC, as Pennebaker mentions it, are that “word counting programs are also remarkably stupid. They can’t detect irony or sarcasm and are singularly lacking in a sense of humor. [...] They fail to capture the context of language. LIWC like almost all word-counting systems, makes lots of errors. Sometimes it classifies correctly and sometimes it doesn’t.” (Pennebaker, 2011, p. 8). With that in mind however, Pennebaker “validates” his work by saying that his Lab has run enough studies that they can say statistically LIWC is usually correct, and the more they use it and the larger the sample being analyzed the more “precise” the analysis is.

Pennebaker has utilized his program to analyze many different texts, from poem of dead poets, to the public speeches of Political figures and has written extensive articles as to what the word choices they make in their everyday speech truly shows the people a glimpse into that person’s life, health and behavioral state. Pennebaker’s work and program served as a base for the study at hand by providing background knowledge and comparison data, as well as psychological analysis to words. Pennebaker is also the only known individual who has created a program that can analyze language in such depth. However, as a Linguist, this researcher knows that there are problems with this validation and as Pennebaker mentioned problems with just

letting the program analyze the language data without human interference. An example of why non-mechanical analysis based on linguistic substance is necessary is described in section 4.2.

2.3 Language and Treatment

The Language and treatment sector deals with how language is used in the treatment area and the importance that language has been given when it deals to implementing treatment to others or in day-to-day speech with the patient/client. The study by Celia Berdes and John Eckert in “The Language of Caring: Nurse’s Aides’ Use of Family Metaphors Conveys Affective Care” looked at how nurse’s aides’ would provide good care without letting racial abuse interfere. In the experiment, they found that the “aides used metaphors associated with family, relationships, and attachment to describe their affective care of residents.” (Berdes & Eckert, 2007) Throughout the experiment, when the aides would be asked to say how their care was effective or not, they would compare how they would apply their care to that of how the families would treat the residents. The investigators found that the way the aides expressed their attachment to the residents was through the use of family metaphors. They would name the relationships they had with their residents. It wasn’t just their patient. The researchers found that some of the aides they interviewed would either: 1) treat the residents like the aide's own parents, 2) treat the residents like their babies and 3) treat them like good friends. (Berdes, 2007). The researchers found that by applying these metaphors into describing the quality of the aide’s care for the residents, they were able to better show the quality of their care. (Berdes, 2007). As stated by Berdes & Eckert, “In this context, the nurse’s aides’ use of metaphors of family and attachment is meant to show us that they are choosing to supply caring affect in their work;” (Berdes, 2007, p. 346) This study showed how language can add to any situation in an applied fashion. Although this article did

not deal with drugs or addictions, it talks about the use of metaphors in the real world and how they reveal mental structures and states.

Martha Zúñiga in “Using Metaphors in Therapy: Dichos and Latino Clients”, talks about the incorporation of metaphors into counseling to create a better rapport with clients. She claims that “use of metaphors in treatment enables practitioners to build a culturally sensitive form of treatment” (Zúñiga, 1992). In other words, in this article Zúñiga is trying to make bonds with her clients through the use of Language and linking her state of mind with that of her clients.

According to Zúñiga, a metaphor is the “application of a word or phrase to an object or concept. It does not literally denote in order to suggest comparison with another object or concept” (Zúñiga, 1992). Zúñiga uses metaphors through stories, proverbs, symbols and objects. She continues to state how there are subclasses of metaphors which include: model, parable, fable, allegory and myth. They all can be taken literally by some people. Her goal in applying metaphors to treatment is to get people to take a metaphor literally but at the same time perceive a symbolic meaning subconsciously (Zúñiga, 1992). This article shows how language is being used in the clinical or therapeutic realm, as well as how particular language uses, i.e. *dichos*, can establish a connection between patient and treatment provider, as well as connection of states of mind.

Another indication of the importance of language used during treatment of substance and drug abuse is a letter to the editor of the American Journal of Public Health written by Sarah E. Wakeman MD from the Massachusetts General Hospital. Wakeman titled her letter “Language and Addiction: Choosing Words Wisely”, in which she focuses on the ill choice of words from the journal that in some articles continued to say “drug abuse” instead of “substance use disorder”. (Wakeman, 2013) Wakeman states that “with our language we imply patients are

inflicting the morbidity of the disease on them and are thus undeserving of care.” (Wakeman, 2013) Wakeman continues with saying that patients tend to be hesitant to get care because of the stigma that is attached to their condition of drug/substance addiction. She thinks that in order to provide better care, practitioners everywhere need to be mindful of the language they use to talk to patients or refer to addictions. (Wakeman, 2013) She states that “even highly trained mental health professionals are more likely to think that a patient is personally culpable and deserving of punitive measures when they hear the patient is described as a "substance abuser” as compared with having a “substance use disorder.” (Wakeman, 2013). This article shows, from a doctor’s point of view, how language can affect perception; (and perception may influence further actions) from both the point of view of a caregiver to that of a patient.

The studies that were focused on in the sections “Language and State of Mind” and “Language and treatment” showed us three different approaches to language analysis. These approaches were: J. Pennebaker’s Linguistic Inquiry and Word Count (LIWC), concordance program which for this study the freeware version of AntConc will be used, and finally the Metaphor implementation to counseling that M. Zuñiga brought to light. LIWC and AntConc analyze language as a whole and Zuñiga’s approach examines what language can be used for. Different approaches for analysis of language gives this study more vantage points on a specific set of data. However, this study wanted to prove the efficacy of language analysis tools without manual (human) interference and therefore, LIWC and AntConc were chosen instead of M. Zuñigas method of analysis.

2.4 Limitations of Previous Studies

Although there are many studies on the language use of people and what it may tell the reader, none have explored or tried to explore if there are any differences through language use between Peer and Professional substance abuse counselors. Comparison of client/counselor rapport and preference exist (Lo Sciuto, Aiken, Ausetts, & Brown, 1984), but do not talk about what could be behind these preferences of the client/patient. Due to these discrepancies, this study sought to, through the use of language analysis tools, compare and contrast the language differences that Peer and Professional substance use counselors could have.

Chapter 3: The Present Study

3.0 Introduction

Research exists in study of language in itself and the connection it has with the mind. Studies even extend out to the analysis of language of certain types of addicts (Pennebaker, 2011), however, nowhere has the language of the actual counselors of substance use treatment been analyzed. This study is the first of its kind where it is just analyzing the language differences between Peer counselors (counselors with previous substance use addiction) and Professional counselors (counselors with no previous substance use addiction) in hopes to serve as a starting point for further examination with a larger sample set to see how possibly in the future it could be possible to bridge whatever is segregating one group from the other in the eyes of the patient or client. The advantages of this study therefore are: 1) having two defined groups for comparison; 2) a defined data set; and 3) a defined empirical methodology and not speculative or anecdotal.

3.1 Research Questions

This study sought to find out if through the use of language analysis tools, if detecting a difference in language use between peer and professional counselors was possible? If it was possible, what could these differences of language tell us about the state of mind of each group of people, or the individual in itself? Would this allow further applying linguistics and its methods to detecting state of mind differences through language use? This study wanted to explore this realm that had not previously been looked into and see what could be yielded through the qualitative analysis of a convenience sample of Directors or heads of drug and

substance use rehab centers. This study is an exploratory study to test the effectiveness of language analysis tools for natural speech analysis.

3.2 Sample and Data Collection

This study is a secondary analysis of data originally collected for the Use of “Juramentos” as a substance abuse social control mechanism in the region of El Paso, TX and Las Cruces, NM area under the Vulnerability Issues in Drug Abuse institutional grant¹ conducted by Dr. Mary Cuadrado. In the Juramento study, 13 directors/ heads of clinics or addiction centers (out of 28 centers contacted) were interviewed². The interviews were recorded using an iPad and the program Super Note. After the data were collected, they were transcribed in Word using Express Scribe and a transcription pedal. Each interview was labeled according to the number the treatment center had on the list of interviews and that provided the center anonymity. For example, the first center interviewed would have been assigned the label T1, meaning treatment center 1. For the purposes of this thesis interviews were divided in terms of Peer and Professional and if there was more than one subject per interview the interview was divided by adding a letter to the name to designate subject 1 or subject 2 or subject 3. For example, if T1 had two subjects speaking there would be file T1a and T1b. The sample of this thesis consisted of a total of 16 subjects; of which 7 were peer counselors and 9 were professional counselors.

¹ DHHS/NIH/National Institute on Drug Abuse (Grant Number: 5R24DA029989-03), Project Title: UTEP DIDARP: Vulnerability Issues in Drug Abuse (VIDA).

² For more Information regarding the Study on Receptivity of Juramentos, Contact Dr. Mary Cuadrado at the University of Texas at El Paso.

3.3 Definitions and Distinctions

For the purpose of this thesis, the category of *professional* will be given to those who during the interview did not say that they had once had an addiction of any sort and who hold a degree. Likewise, the *peer* category will be given to those that specified that they had recovered from an addiction and were now helping others, regardless of them possibly having crossed over into the “professional” realm by acquisition of a degree or working in a salaried position. Once these definitions were applied, there were a total of 9 professional counselors and 7 peer, yielding a total of 16 interviews. The acronym ProvP was used to discuss the Professional versus Peer counselor difference and studies.

3.4 Process

3.4.1 Reliability Checking

To ensure that the transcriptions were correct, a second researcher did a reliability check of the transcriptions. The second researcher, using the program Express Scribe, Word (with the original transcripts), and a transcription pedal listened to all 13 interviews (16 subjects) with the track changes mode on, and changed those words which they found to be different. There were very few differences in the transcriptions. Figure 1 shows a sample of one of the original transcripts in comparison to the reliability transcript, to see what some of these differences were. The differences that arose in the transcripts were thoroughly scrutinized to ensure the most accurate transcription was used for analysis.³

³ Both versions of the transcripts can be viewed if need be. Please contact myself or Dr. Mary Cuadrado for permission to view the transcriptions.

3.4.1.1 Issues with Recordings/Transcripts

Issues that were encountered with the recordings and then translated into the transcriptions are as follows. Due to the fact that most of the places visited were running /active clinics/centers, the issue of sound quality existed. Many times, the location of the interview was in the same hall where the clients/patients would be, or very close by. This caused a lot of background noise to the point that if the subject being interviewed did not project his/her voice adequately, the iPads could not capture the natural speech well. Another issue that was encountered there was a number of 4 interviews where this researcher was not able to attend and therefore had to get the sound file from the PI of the “Juramentos” project. These files however could not be emailed directly and had to be recorded using the playback function of the device and the microphone of another device. This diminished the quality of the audio file for transcription.

Another issue that was encountered in the data was the issue of using Spanish words in their English speech. The fact that the interviews were done in the border region of El Paso/ Las Cruces area, the occurrence of Spanish words in natural English speech was quite predominant in the interviews. Not all instances were an issue since some words have been carried over into the English language. But, since the programs we are using have to use either an English or a Spanish dictionary to analyze the words, it was necessary to make some modifications so that the program would not count certain words; like the Spanish ‘me’ (first person personal pronoun); with an English word that shares the same spelling.

3.4.2 Language Analysis Programs

Each language analysis program used in this study had its own set of configurations that had to be met in order to properly analyze the data. In the following sections, this study will discuss each program and the specific configurations and necessary modifications that had to be done to the data to ensure the best analysis possible with the language analysis program used.

3.4.2.1 AntConc

AntConc (AC) is a shareware version of MonoConc Pro that was created by Dr. Laurence Anthony, a professor at Waseda University in Japan. This program is “an easy-to-use freeware concordance program for Windows (98/Me/2000/NT/XP/7), Macintosh OS X, and Linux. It was originally developed for use by students in the classroom, but also serves as a comprehensive text analysis tool kit for researchers. AntConc is written completely in the Perl 5.8 programming language using ActiveState's excellent Komodo development environment and continues to be developed through feedback from users around the world. AntConc contains the following tools: Concordance, Concordance Plot, File View, Clusters, N-Grams (part of Word Clusters), Collocates, Word List, and Keyword List.” (Anthony).

Although the most common use for AC is for corpus studies, its usefulness for this study became evident from the start. As the website states, this program can give a frequency list of the words that are used in the data inputted, listing from most frequent to least frequent. In addition to providing that list, it gives you collocates that appear most frequently around the words. For the purpose of this study, it was able to give us a word count which could then be calculated in percentages per interview to see how the percentages of certain words or all words used add up. This calculation was necessary to see how each individual file ranked in the words most

frequently used by the group. But, before the program could give any list, some preparation had to be done for the program to work effectively.

Some of the preparations that had to be done to the data were as follows. The very first thing was to change the format from the .docx format to a plain text format so that the program could read it. After that was done each file had to be purged of unnecessary information. Some of that unnecessary information was: the interviewer's speech, the markers for where the subject began speaking; marked by an 'S', and lastly, was the issue of words that were in Spanish instead of English. Lastly, verification of the words, used in the interview, conforming to what the program looks for. Now, given that AntConc is a freeware program, there are multiple tweaks that can be done to the program to fit the corpus under analysis. Some of these tweaks are things like making your own word list of what you want to look for specifically in the text and /or a specific existing corpus you want to compare your corpus to. A feature that seemed helpful for this study at first, was the Lema-list feature. The Lema-list feature as described by the AntConc website is a list created instead of a word-list and it allows the researcher to look for words that are variations of a word. For instance, a lema-list of the word "be" would look like "be->is, are" and it would look for those variations within the text. (Anthony) After those preparations are done, AC allows you to load all the files sought to analyze at one time. This is equal to one corpus. This thesis loaded three corpora, being: corpus 1= all interviews (encompassing of peer and professional), corpus 2= peer counselor interviews and lastly corpus 3= professional counselor interviews.

Once the decision was made concerning how the interviews were to be divided for analysis, they were submitted for creation of word lists. Again, we had a total of 3 word lists, one per corpora, so that later the analyzing of these lists could be compared. As mentioned earlier,

when the word lists are produced along with the frequency that occur in the corpus, it is possible to also look for the words that occur the most around these main words. For example, let us say the word ‘you’ occurred 435 times in corpus 1, it would tell us the frequency and then if we wanted to find the concordance we could choose concordance and it would tell us what other words frequently occur with the word ‘you’.

3.4.2.2 Linguistic Inquiry and Word Count

The Linguistic Inquiry and Word Count Program, or LIWC for short, is “a text analysis software program designed by James W. Pennebaker, Roger J. Booth, and Martha E. Francis. LIWC calculates the degree to which people use different categories of words across a wide array of texts, including emails, speeches, poems, or transcribed daily speech. With a click of a button, you can determine the degree any text uses positive or negative emotions, self-references, causal words, and 70 other language dimensions. The LIWC program can analyze hundreds of standard ASCII text files or Microsoft Word documents in seconds. The LIWC2007 program also allows you to build your own dictionaries to analyze dimensions of language specifically relevant to your interests.” (LIWC Inc.) Unlike AntConc, LIWC does not just count “raw” words, but it groups them into categories. The categories into which the words are grouped, can be either grammatical (or syntactical) and other can be semantic categories, obvious ones versus not so obvious. Dr. Pennebaker in his findings and writings has claimed to have validated the semantic categories that he has created for the program (as can be seen in more detail in chapter 2), although there is somewhat of a controversy over the groupings and the efficiency. Also, the categories formed for the LIWC program consist of many words that may be found multiple times in the dictionary for each separate category (for a glimpse at the dictionary please refer to

table 1 in appendix). For example the word “ate” would fit into the category ‘verb’ and ‘past tense’. For the purpose of this study, however, only those categories that produced a significant difference (as explained in 4.2.2) were addressed due to time constraints and research question asked.

3.5 Order

The order this study applied the language analysis tools was for the purpose of practicality. Therefore, AntConc was applied first to give us word count and frequency of words used, and applied intermittently during the LIWC analysis to see individual counts for LIWC categories and to help analyze LIWC findings in context of the interview. AntConc gives us the capacity to be able to search for a particular word in question and see how it was used within the interview. For example, let us say that the word “you” wanted to be seen in terms of usage or meaning of use in the interview, one would proceed to search for the word “you” then see the output in terms of collocates and it shows you the sentence in which the word in question was found. Both however worked hand in hand and gave a deeper view into the analysis, one (AntConc) in terms of a more basic analysis and second (LIWC) in a more in depth analysis of every single word used in each interview.

3.6 Merging Changes in Transcripts

From the thirteen interviews (16 subjects) there was an average of 7.25 mistakes in the transcriptions of the subjects' speech per interview in the professional interviews, and 5.8 mistakes in the transcriptions of the subject's speech in the peer interviews. The word count of all the interviews equaled 36,513. A detailed table of the number of errors per each interview is

included in the Appendix of this thesis (see table 14). Some of the common mistakes that were found throughout the transcripts were different fillers, i.e. instead of transcribing "um" they would transcribe "and". But also in some instances there were completely different words that had been heard between the two researchers. In those instances, this researcher went back and listened to that specific section of the transcript to double check what the second researcher had heard and if this researcher did not agree with what the second researcher heard, a third researcher was brought in to check the discrepancies and rule which interpretation was correct.

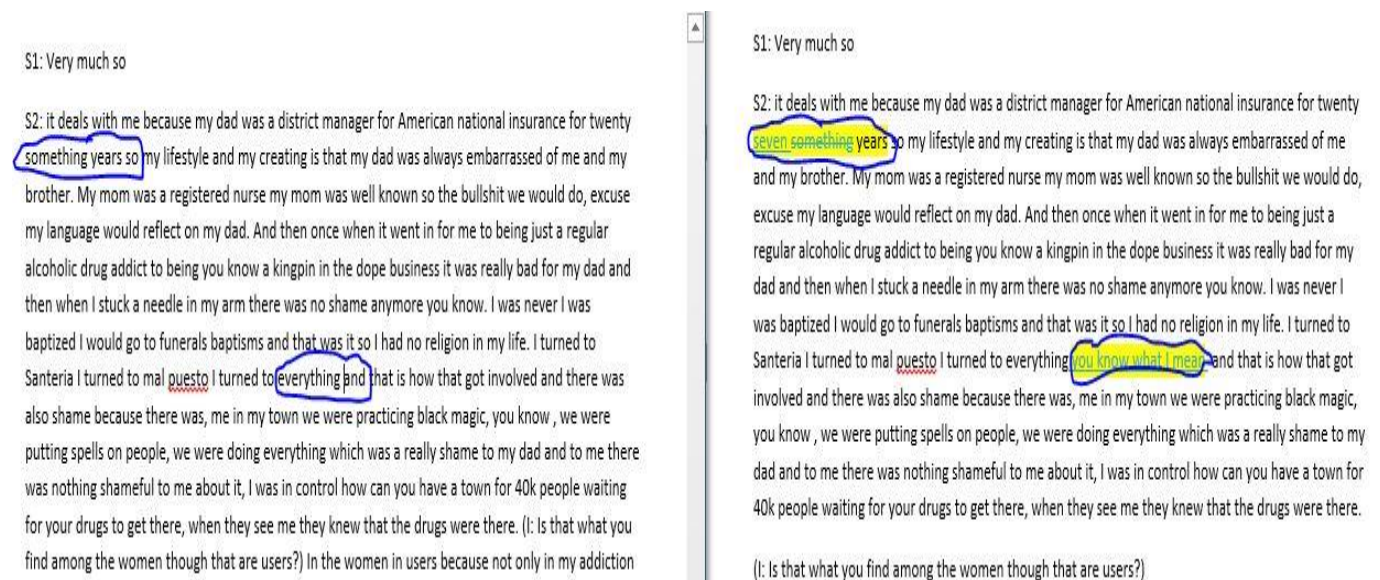


Fig. 1.1 Left Original Transcript, Right Reliability Checked Transcript

3.7 Preparing Transcripts for Analysis

Since this study used natural speech, and each language analysis program requires the transcript to be molded to the specifications of the program, some changes had to be made to the transcriptions. Many of them were just making conjunctions, placing brackets around fillers, etc. so that the program would read the transcript correctly. An important change for AntConc was changing the format from a Word format to a plain text format. Also, the transcripts were of live

interviews and originally included the Interviewers speech along with the subjects. Since we are not analyzing discourse, nor the language of the interviewer, each interview transcript was edited to remove the interviewer's dialogue. Also, if an interview had more than one subject's speech, the interview was then divided into different files so that each subject could be analyzed independently.

3.8 Sample

3.8.1 Description of the Sample

For the purpose of this study, secondary data analysis of a convenience sample of two groups of counselors found in the El Paso/ Las Cruces area was done. The individuals were Directors or representatives of drug and alcohol treatment programs who volunteered to participate in the study. The groups were divided into peer counselors of drug and substance abuse programs and professional counselors. In order to make this distinction, and for the purpose of this study, those that mentioned even once that they had at one point been an addict would be placed in the peer group regardless of educational background, and those that never mentioned they had been addicts were in the professional group. Since the sample was a convenience sample, as well as a qualitative sample the size was conservative. The gender of the sample was predominantly male: 12 out of 16 subjects were male. The ethnicity of the sample was predominantly Latino: 13 out of 16 subjects were Latino. Since age was not requested from any of the subjects, this study cannot comment on age differences for certain.

3.8.2 Sample Excerpts from Interviews

In the following two sections the inclusion of an excerpt from each pool of interviews will be shown. This is so to give a quick glance at some of the characteristics that stood out most during the interviews process.

3.8.2.1 Peer Sample

T10b excerpt:

“S2: When I walked through theses doors or walked through prison I had nothing. My family had given up on me. The only reason I had some contact with my family is because my mom, I called my mom and told her I have to make amends with you and my grandmother. So they brought by both of them and that’s part of growing and accepting responsibility accountability, you know what I mean. So that’s what it was for me and then he said well you know what your kids want to see you, you know.”

3.8.2.2 Pro Sample

T1 excerpt:

“S: Well for everybody that does come in we have the psycho-education portion of the treatment first. Which we do have towards the family system, to a little bit of a belief rational challenging and kind of more of a natural consequence exploration. Kinda like you know you happen to do this but you’re about to lose your kids, type of thing, so more of a rational belief challenging part of the treatment and that’s how the beginning, we of course follow parts of the stages of change where you know the person might be more in the pre contemplation stage, but I I couldn’t say that we have a cookie cutter recipe for everything but it has to confront them”

Chapter 4: Results

4.0 Introduction

Just how the analysis portion of this study was divided by program used, so were the results to be able to give a clear view of what was generated by what and then later how both programs helped this study reach its conclusions about the question at hand. This study begins, once again, with the simpler of the two programs AntConc, and then continues with LIWC. A third program was included in the analysis stage and that was the use of SPSS with the results from LIWC to help narrow down this study's analysis of the 64 categories yielded from LIWC to the most significant.

4.1 AntConc

By starting with the most general of the methods, AntConc, we were able to see what output we got within the level of words before applying LIWC and looking at the semantic categories. As previously mentioned, each program required its own preparing of the original transcription of natural speech in order to fit the norm required for the program to function as desired.

4.1.1 Results

AntConc is the simpler program of the two. The purpose of it, for this study, was to generate frequency word lists. As mentioned in chapter 3, three lists were generated. These lists stated the number of times the word was produced in the corpus used and the percentage of occurrence. Since AntConc produces a count of every single word, each list is a bit lengthy. For the purpose of this thesis however, the focus was on those words which yielded a higher

frequency of use as well as percentage. In order to differentiate between lists, each one was given a name: 1) ProvP Ant, for the list yielding the whole corpus including Professional and Peer interviews, 2) Pro Ant, for the list yielding only the Professional interview frequency word list, and finally 3) P Ant, for the list yielding only the Peer interview frequency word list. The program yielded the number of words per corpus in the results as well as collocates that most frequently occur before or after a word. For list 1 (ProvP Ant), the total number of words was 34,835. The most frequent words were: 1) 'the' with 1203 occurrences, 2) 'and' with 1135, 3) 'to' with 1123, 4) 'I' with 1031, 5) 'that' with 955, 6) 'of' with 760 and 7) 'we' with 758, just to name a few. Refer to Table 2 for results of the ProvP Ant list up to rank 25. List two (Pro Ant) had a total of 17,919 words. The most frequent words were: 1) 'the' with 670 occurrences, 2) 'to' with 598, 3) 'and' with 519, 4) 'that' with 516, 5) 'I' with 450, 6) 'we' with 425 and 7) 'of' with 415, just to name a few. Refer to Table 3 for results of the Pro Ant List. For list three (P Ant), the total number of words was 16,916. Of those words, the most frequent were: 1) 'and' with 616 frequency hits, 2) 'I' with 573, 3) 'the' with 530, 4) 'to' with 514, 5) 'you' with 413, 6) 'that' with 384, and 7) 'a' with 342, just to name a few. For a list of the top 24 most frequent words in the 'P Ant' list refer to Table 4.

4.1.2 Discussion of AntConc Findings

Although the word lists created by the program AntConc can serve to show words used repeatedly in each List of interviews, that is all it truly gives us: a list. Unless a pre-determined word bank is established, the list generated can become overwhelming. The results from the lists were unrevealing since as mentioned in the results sections the most frequent words in this corpus, and any other corpus are going to be words like: articles, conjunctions, etc. Therefore, as

previously mentioned, the use of AntConc was a jumping board into a program that goes beyond generating a list of words with the frequency and collocates.

4.2 LIWC

4.2.1 Results

As mentioned in the methodology chapter, LIWC has the capacity to generate and place words from a bank into 60+ grammatical and semantic categories. This procedure was applied to the data, which was divided into two groups, or lists, instead of the previous three lists used with AntConc. LIWC yielded a report with 64 categories, in which it gave a percentage of how many times a certain type of word or word itself was used in each interview in comparison to the dictionary that is assigned to each category in the program⁴.(Pennebaker, 2011, p. 8) This created a table (shown in Tables 5 and 6) of 66 columns and 10 rows. However, those 66 categories were then narrowed down to 64 for SPSS since the first two columns just gave information of peer/pro and file name. The output yielded directly from LIWC gave this study a number of percentages per category and per file which, as stated previously, is calculated by how many times a category from the predetermined dictionary for such category is found in the file being analyzed. Within the group there were immediate numbers that stood out per category and specific categories that had higher percentages within them as well. When looking at the LIWC output all 64 categories were looked at to attempt to find a pattern of use among the subjects. When a number and category stood out, i.e. the percentage for one subject was double any of the

⁴ Each category in LIWC can contain up to 700 words in the dictionary for that category which was predetermined by the creators of the program. For further exploration into the dictionaries refer to chapter 3 and table 1.

others, that category and file was marked by an asterisk for further analysis. This occurred a few times within both groups, Peer and Professionals.⁵

After highlighting and analyzing subjects' differences within their group, this study compared groups to each other by creating a mean of each category for each group (instead of individual). This process revealed where each group differed from the other and where they were similar. However, to statistically see the significant difference in all the categories of LIWC between Peer and Professional counselors, the program SPSS was used. It must be noted however, that during an analysis of word count and pronoun count through the use of AntConc it was seen that a file, T5c to be exact, was an outlier due to its word count size. Unlike the other interviews in our sample, T5c consisted only of 233 words total, where all the other interviews had over 1000 words each. Therefore, it was decided to exclude the file of T5c from the data analysis changing our sample size from 16 to 15 individuals.

4.2.2 SPSS

To avoid subjectively deciding which categories of words were used sufficiently different between peer and pro counselors, the data was inputted into SPSS, and independent sample T-test was used. Table 7 shows the results for the significant categories in a T-Test of both the peer and professional statistics. Test result significance was determined by using the 2-tail of .10 or less.

Out of the 64 categories that were analyzed 49 of them showed no significant differences and 15 showed significant differences amongst the two groups. Of those 15, Peers averaged

⁵ For an example of this, please go to Appendix table 5 and 6 which shows the full table of LIWC results for both Peer and Professional Counselors respectively.

higher use in 9 categories, which included: the linguistic processes ‘you’, ‘he/she’ and ‘swear words’ categories. In the Psychological processes category, Peers had a significant difference in the area of ‘humans’, ‘perception’, ‘hear’, ‘biological processes’, ‘body’ and ‘sexual’. Professionals averaged a higher use in 6 out of those 15 categories with significant differences which included: Linguistic processes ‘they’, ‘articles’ and ‘future’; Psychological Processes ‘discrepancy’ and ‘exclusion’; and finally, Personal Concerns ‘work’. It must be noted once again that these categories and groupings of words under certain categories were predetermined by Dr, Pennebaker and students using judgment test of words. This could have some linguistic problems if semantic properties of certain words were not accounted for and grouped together anyway (please see section 2.2.1 and chapter 3 for further explanation of LIWC methodology).

Due to its vast possibility of meaning the results were divided into sections for discussion: Pronouns, Psychological Processes, Linguistic processes, and all other Categories. These divisions were chosen based on what SPSS gave as an output, as well as how the program LIWC splits up or groups the categories together.

It was also decided to give Pronouns their own section instead of including them under Linguistic Processes since pronouns are known to carry so much more information about the person speaking and are more specific than the other categories created by LIWC. It must be noted, in the dictionaries pre-determined by LIWC the pronoun categories include more than just your basic pronoun, i.e. “you, he/she, they, I , we” but also include other variants of those words for example: “their, your, our, me etc.” Therefore, the raw numbers yielded by LIWC under this category would also include the variants of that word.

Table 7 T-Test Significance of Data

File	Category	Sig 2-tail	Mean Peer	Mean Pro
V8	You	0.029	3.0617	1.5278
V9	She/He	0.021	0.7433	0.3022
V10	They	0.026	2.65	4.0978
V12	Article	0.091	5.5217	6.3978
V17	Future	0.081	0.7283	1.1544
V24	Swear Words	0.051	0.0267	0.0033
V28	Human	0.019	1.8817	1.2156
V38	Discrepancy	0.083	1.7167	2.3633
V43	Exclusion	0.082	2.9267	3.7567
V44	Perception	0.105	1.6683	1.2911
V46	Hear	0.015	0.8317	0.4767
V48	Biological Process	0.101	1.6717	1.3011
V49	Body	0.026	0.2	0.0722
V51	Sexual	0.048	0.1233	0.0633
V57	Work	0.1	1.6067	2.4789

4.2.2.1 Pronouns

As previously noted, the output for LIWC yields a total of 64 categories. After running the LIWC data through SPSS the categories were narrowed down to 15 categories total. Of those 15, three of them were pronouns. The pronouns that showed significant differences between the groups were: You, They, He and She as exemplified in table 8 and 9. The numbers shown in table 8 and 9 were extracted from the raw LIWC data to see how each individual interview scored in the category that showed significance. Also, a mean was provided, which was used to see how each individual file compared according to the mean of the interviews in their group (peer or professional). It must be noted that those individual files that scored above and below

the mean of the group will be discussed further to explore as to why those differences could have occurred. Also, it must be noted that these categories include more than just the standard “you, he, she, they” pronouns, but include the variants as well, such as he'll, he'd in addition to /he/, etc. The matter exists that will be touched upon further about the pronoun “you” and its variations of non-real second person pronoun “you”, i.e. you all, you (indirect) that this study will talk about in more detail in the section to come.

Table 8 Professional Pronoun use and Means

LIWC Pronoun Pro Use				
File	Word Ct	You	They	He/she
all files	17919	13.75	36.88	2.72
File	Word Ct.	You %	They %	He/She %
T1	2547	1.29	3.61	0.27
T3	1849	1.56	5.5	0.29
T4	1661	1.95	4.1	0.13
T5a	1429	1.64	3.66	0.07
T6	2322	1.04	1.73	0.91
T11	3526	4.04	5	0.24
T12	1276	0.32	4.62	0.24
T13	1130	0.94	4.41	0.47
T14	2179	0.97	4.25	0.1
Total Mean		1.53	4.10	0.30

**note: percentages were generated through LIWC comparing text to an extensive dictionary per category of possible words and then averaged of the times used through interview.*

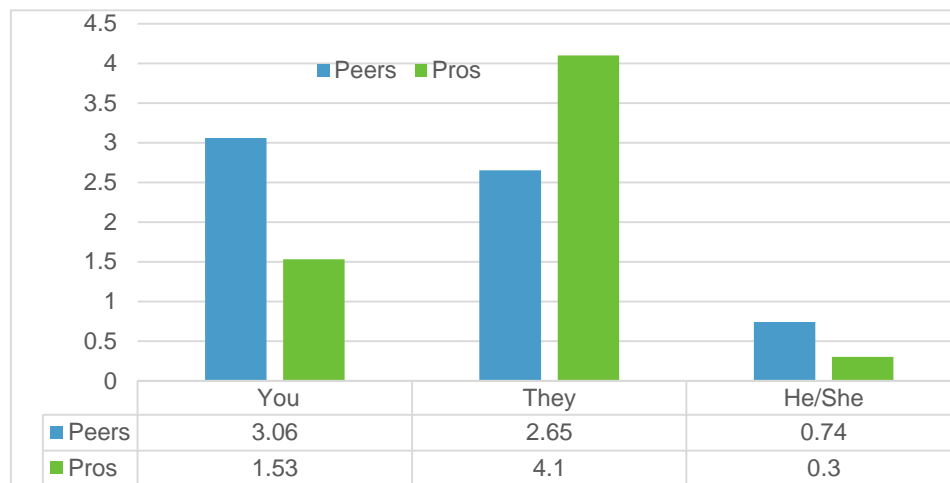
Table 9 Peer Pronoun use and Means

	LIWC Pronoun Peer Use			
File	Word Ct.	You	They	He/She
ALL Files	16683	18.37	15.9	4.46
File	Word ct	You %	They %	He/she %
T5b	1806	4.61	3.55	0.41
T7	3067	1.95	4.13	0.64
T8	1628	1.99	3.26	0.27
T9	4286	1.66	1.81	0.78
T10a	2476	4.65	1.45	1
T10b	3420	3.51	1.7	1.36
Total Mean		3.06	2.65	0.74

**note: percentages were generated through LIWC comparing text to an extensive dictionary per category of possible words and then averaged of the times used through interview.*

As a group, Peers ranked higher in the use of ‘You’ and ‘he/she’ than the Professional counselors. Professionals ranked higher in the use of ‘They’ than Peers, as can be seen in Graph 1. Table 8 and 9 show the exact percentage of use per individual file per pronoun category. A mean calculation was derived to be able to see how each file ranked compared to the mean of all files combined. By comparison of mean to individual file we see a few files that ranked noticeably higher or lower than the mean of all files. In terms of the pronoun “you” the files that ranked way above or below the mean for the Professionals were files T11 with a percentage of 4.04 and T12 with a percentage of .32. Those two files out of the nine acted differently than the rest. For the Peer counselors the files that acted differently than the majority were that of T5b with a percentage of 4.61 and T10a with a percentage of 4.65. The following analysis could help determine what could have made these files act differently, or if there is a different reason as to why these files ranked higher in the use of “you” than their counterparts.

Graph 1 Pronoun Use Difference between Peer and Pros



According to Pennebaker, pronouns have a deeper meaning into the mind, and can give great insight as to the type of person that is speaking, or writing. (Pennebaker, 2011, pp. 1-17) The studies of Kacewicz *et al.* 2013 and Pennebaker 2011 tell us how those in higher social status rank higher in the use of “You” and “We”, where as those that are lower in social status rank higher in the use of “I”⁶ Let us now analyze what the regular standard use of ‘you’ as second person pronoun would mean. According to Pennebaker, the use of “you” pronouns can exemplify not only higher social status but also: anger, focusing on their audience/others, being deceptive and “you-words as the equivalent of pointing your finger at the other person while talking” (Pennebaker, 2011, pp. 107,174) With that definition in mind, one thing that could account for the use of “You” in the Peers’ speech is that they presume themselves as higher rank amongst their own group of people. Another possibility is that they are trying to exemplify their authority or higher rank status during the interview and that is why subconsciously they are using ‘You’ at higher rates. If the way the Peers used “you” was in the sense of deception, a possible

⁶ What is meant by social status does not necessarily correlate with economic status, but more of submissiveness versus power, boss versus employer and those types of relationships

explanation for that could be that they are again making up for, or talking themselves, or their program “up” to rise to the status of the interviewer as well. If, however, the use of ‘You’ for our peer group where used as the ‘pointing of finger’ sense, it would make someone wonder as to why addicts would “prefer” or open up more to a person that is giving off that vibe onto the addict.

However, although these explanations are plausible with our analysis of “you” use by Peer counselors, it is imperative to remember that linguistic analysis must be done manually, in addition to any computer based analysis, to ensure that proper analysis is being performed. A good reason for this could be the fact that pronouns, for example, such as “You”, when just looked at alone could be taken to mean literally the second person. However, one must also account for using the pronoun ‘you’ in instances where it is not referring to just one individual but many or none at all and rather as a figure of speech and or filler word, i.e. you all, you know. Pennebaker (2011) does not account for this distinction of different types of “you” in the speech but just analyzes different meanings of just traditional second person pronoun “you” and assumes that all “you” are the same for his psychological interpretations of the second person pronoun. For linguistic purposes however, since “you” is not always second person pronoun, it can be imperative to manually grade, or label or analyze the transcripts.

In regards to the multiple types of “you” that could have an effect on what is really meant by the speaker, it was necessary to go back to the interviews (with the help of AntConc) and search for where the word occurred in the text and see how the word was used semantically. This process revealed the following: Peers still almost doubled in use of “you” in the specific sense and the general sense compared to Pros. However, both Peer and Professional counselors had a vast amount of use of “you” in the sense of the filler “you know”. This difference was not caught by

the program LIWC since this distinction cannot be classified yet by computer programs. It took a manual reading and interpretation of the script to decipher which “you” was being used. This is something to consider as to where else this may have occurred. Statistically however, Peers still had a total of 254 instances of the use of “you” where Pros had a total of 156 (in the specific second person pronoun use). This can be seen in Table 10 and 11. This still coincides with our analysis of the fact that Peer counselors use “you” (specific second person pronoun) more than professional counselors. The proper separation of “you” from its other forms confirmed that the files that were already acting differently from its counterparts were still doing so in terms of using the specific second person pronoun “you”, as can be also seen by table 10 and 11. This would lead us back to the definition that Pennebaker gives for the use of “you” previously mentioned. If this definition and interpretation offered by Pennebaker were to be applied to the new segregated correct analysis of “you”, this study can attempt answering the question of what does it mean that peers still prefer the use of “you” over most professional counselors (since there are those two files were the professional counselors did use a vast amount of the second personal pronoun “you”).

Table 10 Different types of “You” Peer counselors

	<i>You (specific)</i>	<i>“you know”</i>	<i>You (general)</i>
<i>T5b</i>	64	13	5
<i>T7</i>	32	20	6
<i>T8</i>	18	13	6
<i>T9</i>	12	53	3
<i>T10a</i>	67	11	15
<i>T10b</i>	61	39	6
<i>Total Peer Use</i>	254	149	41

Table 11 Different types of “You” Pro counselors

	<i>You (specific)</i>	<i>“You Know”</i>	<i>You (general)</i>
<i>T1</i>	12	15	3
<i>T3</i>	10	16	0
<i>T4</i>	13	16	1
<i>T5a</i>	5	18	1
<i>T6</i>	14	3	2
<i>T11</i>	86	44	1
<i>T12</i>	1	3	0
<i>T13</i>	3	6	0
<i>T14</i>	12	7	0
<i>Total:</i>	156	128	8

Studies that have evaluated the relationship or rapport between patient/client and counselor have shown that those in recovery open up more and have a closer relationship to those counselors that were ex-addicts themselves (Lo Sciuto, Aiken, Ausetts, & Brown, 1984), but the question still remains as to why if both types of counselors, peer and pro are helping them, or trying to help them reach the same goal. The use of ‘you’ does not help explain this preference, if in fact it is being used in the almost “judgmental” sense. However, one could argue that the reason why the peer counselor feels the earned right to be “judgmental” and use at high frequencies the pronoun “You” could be because he or she feels they have earned that right. Or as a peer counselor said during one of our interviews “no offense but I did not learn what I know from books” and because they have overcome the addiction feel it is in their right to tell the patient/ client “how it is”. This would then fully explain as to why peers can “get away with” or why addicts accept the “judgmental” use of “you” from a peer counselor to them.

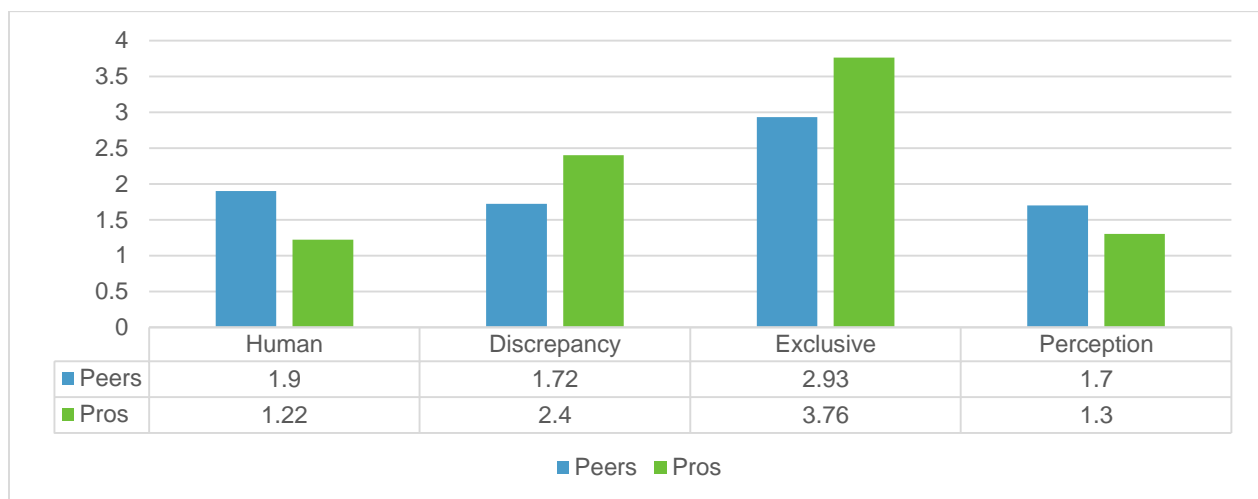
Moving on to the last group of pronouns, the third-person pronouns are known to be detachment pronouns, meaning that whomever is using them is viewing the situation from a detached point of view. Both peer and professional counselors had one of the categories in which they scored higher. Peers ranked higher in the use of “he/she” and Professionals ranked higher in the use of “they”. Although both categories are said to be detachment pronouns, the level of detachment from the situation seems more pronounced with the category pronoun of “they” versus “he/she”. When an individual uses the pronoun “they” it is strongly separating the speaker from a group of people that he or she does not belong to. Yet, the use of he or she, is just separating the speaker from an individual. Although both separate the speaker, one pronoun separates the speaker more so than the other. This subtle distinction, could be the key as to why clients/patients of the peer and professional counselors would have a closer relationship with one over the other. It would also explain the fact that studies have noted how peer counselors are too “close” to the situation at hand and may try to be detached but cannot fully because of their own experience with addictions (Doukas & Cullen, 2010).

4.2.2.2 Psychological Processes

Aside from the pronoun category emerging, other categories that LIWC yielded was as LIWC labels it Psychological Processes, biological processes and social concerns. This category was made up of many subcategories like: cognitive mechanism, discrepancy, exclusion, perception, body, hear, biological process and work, to name a few. Again it must be noted and refer back to section 2.2.1 on how these categories were created by Pennebaker and his students as well to refer to table 1 to see what kind of words are included in each category.

Some of the things that stood out for the data were that Professional and the Peer Counselors differentiated in some of the categories that fell under this umbrella category. The peer counselors tended to use more words that described bodily functions, or senses as well as social labels. On the other hand, the professional counselors would use more words dealing with the mind and cognitive processes. (See Graph 2).

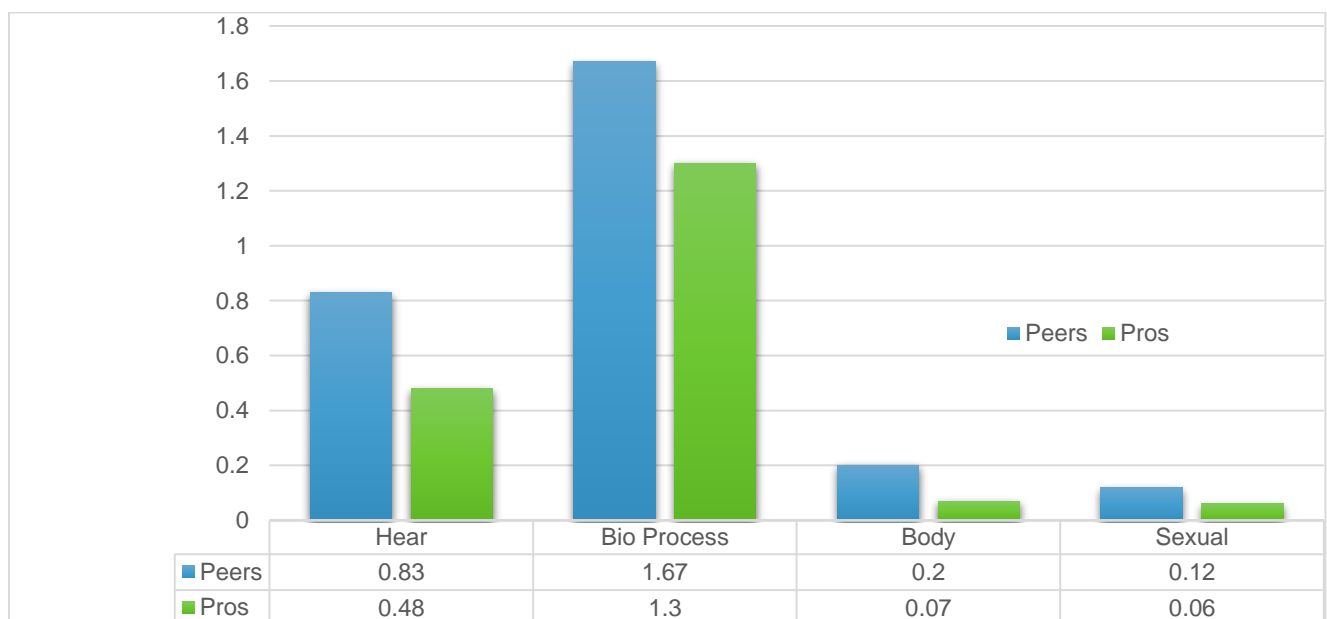
Graph 2 Psychological Process word use difference between Peer and Pros



The categories that fell under this umbrella and had a significant difference were: Human, discrepancy, exclusive, perception, hear, biological process, body, sexual and work. The categories in which professional counselors ranked higher was that of discrepancy, exclusion, perception and work. Peers ranked higher in the categories human, hear, biological process, body and sexual. The category of ‘human’ encompasses words like “boy, baby, Adult”. This shows that mentally Peers are speaking and using more labels of people. The category Discrepancy includes words such as “should, would, could”, thus relaying uncertainty. The category ‘Exclusion’ has words such as: “but, without, except”. This category is known to be used by

those of analytical thinking amongst other interpretations. Pennebaker states it as “people who work to understand their world”. The category ‘Perception’ includes words such as “observing, heard, and feeling”. The above categories were described as “mind” categories whereas the remainder are considered “body” categories. Category Biological Process includes words like “Eat, blood, pain”. Category Body includes words like “cheek, hands, spit”. And the category Sexual includes words like “love, horny, incest”. In a sense this distinction gave the interpretation of separation of self in the subject versus not. It could be that professional counselors during the interview process separated themselves from the interview and therefore their word usage contained more cognitive words since the experience did not pertain to them. On the other hand, it could be speculated that peer counselors, because the topic was “close to home” per say put themselves into the questions that were asked and therefore used more body and emotion heavy words that represented their own personal experience to the questions being asked. (See Graph 3). This creating a possible issue that Doukas and Cullen had mentioned in their article. (Doukas & Cullen, 2010)

Graph 3 Psychological Processes Peer v. Pro Continued



4.2.2.3 Other Linguistic Categories

The last three categories that showed to have a significant difference according to SPSS in LIWC were that of articles, future tense, and swear words. Pronouns would normally be included under this umbrella category of Linguistic processes, but, since pronouns carry such importance and can reveal more about the speaker than these last three categories, this study decided to give pronouns their own heading and to have discussed them first. The categories of articles, future tense and swear words each can have their own interpretation. As can be seen by table 12 and 13, professional counselors had a significant difference in the category of articles and future tense word usage, and peer counselors had a significant difference in the category of swear words.

Table 12 Professional Linguistic Categories

		Pro Ling Category		
File	Word Ct	Article %	Future %	Swear Words %
all files	17919	57.58	10.39	0.03
T1	2547	6.83	1.37	0
T3	1849	8.29	0.41	0
T4	1661	5.36	1.32	0
T5a	1429	5.9	0.82	0
T6	2322	7.22	0.95	0
T11	3526	5	1.08	0.03
T12	1276	7.46	1.3	0
T13	1130	5.82	1.88	0
T14	2179	5.7	1.26	0
Mean Total Use		6.40	1.15	0.003333333

**note: percentages were generated through LIWC comparing text to an extensive dictionary per category of possible words and then averaged of the times used through interview.*

Table 13 Peer Linguistic Categories

File	Word Ct	Peer Ling Category		
		Article %	Future %	Swear Words %
all files	16683	33.13	4.37	0.16
T5b	1806	5.2	1.12	0
T7	3067	5.61	0.44	0
T8	1628	5.38	1.4	0.07
T9	4286	5.74	0.76	0
T10a	2476	6.27	0.25	0.04
T10b	3420	4.93	0.4	0.05
Mean Total Use		5.52	0.73	0.03

**note: percentages were generated through LIWC comparing text to an extensive dictionary per category of possible words and then averaged of the times used through interview.*

Within each category though as can be seen by the mean use, it is viewed how some subjects were above or below the mean word use of that category. This could account for the difference in linguistic style or other possibilities, like education level of each individual. The differences of some of the subjects will be discussed with each pertaining category discussion.

The use of future in speech, or rather repetitive use of the future tense, relates to worrying about the uncertainty of the future instead of living in the present or worrying about the past. The fact that professionals ranked higher in use of future tense in their speech correlates with this analysis since they are not sure of the outcome of the treatment they provide. It must be noted that this uncertainty is not tied to efficacy of the treatment program offered by the professional counselors, but instead it is more so tied to the outcome relying on the person partaking of the treatment and their conviction to it. Since, the future of the clients partaking of the treatment is unknown, it could be said that focus of the future is prevalent amongst the speech of Professional counselors, according to our data. Now, if the observance of which subject ranked higher or lower than the mean use of the future word category is taken into

account, it must be noted that professionals except for one subject (T3) all ranked average. Meaning they all used the future tense evenly. T3 was the only one that underused the future tense compared to the mean use by all the other professional counselors. This discrepancy could be accounted for a less concern of the outcome because possibly of success rates he or she experienced, or maybe this counselor was more “relaxed” or accepting that the future is out of his or her hands to control. It could also be due to linguistic style that could have been the difference, or at the moment the interview was done his or her focus was not on the future but in the present or past.

The Peer counselors out of the 6 subjects two ranked higher than the mean use of future tense in their group. T5b and T8 ranked higher in mean use of future tense words and T10a ranked lower than the mean use of future tense words. As it can be seen, peers also had use of future which future use explanation could also apply, but, it is believed that the reason why Peers may not have used the future tense as prevalent as professional counselors ties back to the issue that Peer counselors put themselves, and their own experiences into their own treatment and may think “if I was able to do it, so will my client” and therein lie the reason why their uncertainty is not as high.

Articles, and their use are related to analysis and analytical thinking process. According to Pennebaker, people who are analytical use articles at a higher rate. Analytical people are said to be distanced from situations to get a better perspective in the analysis of the situation at hand. Professionals, once again following the pattern of the previous significant categories, ranked higher in the use of articles than their counterpart counselors. In terms of mean use, two of the 9 subjects in the professional category ranked higher than the rest. These were T3 and T12. T11 ranked the lowest of the 9 professional counselors in mean use of articles. This finding, further

strengthens the idea that through speech Professional counselors are portraying a more distanced approach to the situation and the questions that were asked.

Peers, had a higher significance in the use of swear words. Linguistically, swear words are associated with, or tied very closely to emotions. Emotion heavy words, like swear words, can at times really exemplify what is being thought and felt by the speaker. However, it can also be said that the use of swear words can be related to either group identity, i.e. the social group you belong to either accepts or rejects the use of swear words, or subtlety of the individual speaking. Either way, swear words are emotionally and socially linked and are considered by this study to be “body” words. Peers having a higher significance in swear words, correlates with the other findings of the group showing to be emotionally involved through their speech, to the situation at hand and questions that were asked.

4.2.2.4 All other Categories

As previously stated, the program Linguistic Inquiry and Word Count (LIWC) yielded a total of 64 categories for each individual interview that was inputted to be analyzed. Out of those 64 categories, after inputting the data through SPSS to find the significance value and to see which categories had a significant difference between both groups of Professional and Peer counselors, 15 categories showed to have such difference. The other 49 categories, did not have a significant difference between the two groups. It could be speculated that the reason why 49 out of those 64 categories had no significant difference could be because these two groups who are seen as “different” are not as different and they are made out to be. Another possibility could be that if a comparison were done between peer, professional counselors and a regular civilian we would find many similarities as well.

Let us revisit the first possibility however. Could it be that since both of these groups have the same goal to provide the same aid and treatment to others that they have internalized, subconsciously perhaps, into their speech so many things that when they speak, although 15 categories do differentiate them, the other 49 do not. It could be, that although still the patient is seeing them differently because one has had prior experience in addiction personally, and the other has not that both groups of counselors are not that far off from one another and could maybe find a way through their language of bridging the differences.

4.3 Integration

Language has been thought to be the gateway to the mind or thought of a person. Through the use of language analysis tools, this study has been able to analyze the language of two groups of people that share the same vocation, Peer and Professional substance use disorder counselors. Previous studies (Lo Sciuto, Aiken, Ausetts, & Brown, 1984) have shown that patients/clients of treatment centers tend to gravitate towards, and open up to, the counselor that has shared experiences of an addiction, regardless of that person's success rates in clinical treatment. Which in turn shows that somehow the client/patient looks for someone he/she can relate to.

In terms of language, this study has shown that the two groups do have differences, and the differences found tie in to one another, or accentuate the findings even more. This study has found that Professional counselors exhibit through their speech a more reserved and distanced approach to the situation, whereas Peer counselors exhibited a more connected and involved approach. This again does not necessarily correlate to the efficacy of their treatment but to what the counselor themselves are bringing into their treatment approach, consciously or not by way

of speaking. The fact that none of the counselors were present when the other counselors were interviewed, and that each was at a different facility (with exception of two) it was interesting to see how the speech of the Peer and the Professional counselors acted in the same ways as their fellow group members. It is imperative to see that although there were evident differences in their speech, and possibly mind process, there were also similarities. This would give way for further research as to how these differences could be mediated so as to have more patients/clients not favor one group of counselors over another.

Chapter 5: Discussion and Conclusion

5.0 Introduction

The purposes of this study were to: i) initiate study involving analysis of difference between peer and professional counselors, ii) investigate how the language of peer counselors in comparison to professional counselors could reveal a difference in the state of mind of each, and iii) analyze the efficacy of language analysis tools to analyze natural speech without human intervention. Since speech is said to be the window to the mind, or the closest one could get to understand the working of the mind, this study tackled this issue of “difference” through the realm of speech.

This study did secondary data analysis of data collected for a DIDARP funded project which had interviewed directors/representatives at a total of 14 treatment centers in the El Paso/Las Cruces area in regards to receptivity of “Juramentos” incorporation into their treatment programs. The data collected was made up of 16 individuals who gave treatment at these facilities. Of those 16, 9 were what this study referred to as Professional Counselors (counselors who had never had personal experience with addiction, regardless of education level) and 7 were Peer counselors (counselors who had personal experience with addiction, regardless of education level).

The results showed that there was in fact a difference through speech between the two groups: 1) Peers tended to use emotional and “body” words, as well as the pronouns “you” and “he/she” with greater magnitude than professional counselors, thus exhibiting personal involvement into the situation yet with an almost “judgmental” feel and 2) Professional

counselors tended to use words that showed detachment and distance from the situation and group concerned, i.e. “mind” words and the pronoun “they”.

It was also proven that language analysis tools cannot yet analyze natural language without some sort of human intervention to ensure that the language is being analyzed correctly. This was seen very specifically with the program LIWC in terms of analysis of the second person pronoun “you”. To the untrained mind, or someone who is not a linguist, this key point could have gone unnoticed and improper analysis would have occurred. Although, the findings were still very similar that Peers ranked higher than Professionals in the use of standard second person pronoun, it was imperative to have dissected the different “you” that were used in the natural speech.

5.1 Previous Research and Current Findings

The lack of current studies attempting to explain as to why patients/clients of substance use disorder treatment centers prefer to have a peer counselor versus a professional counselor, gave way to this study which sought to explain through language if an apparent difference existed, and if this difference could tell us about the mind process of the individuals speaking, in this instance those individuals being the counselors themselves. This study’s data showed that differences between professional counselors (those that never stated they had an addiction to a substance before, regardless of educational level) and peer counselors (those that stated they had at one point in their life had a substance addiction, regardless of educational level) existed.

While the professional counselors through their language exhibited a distancing off from the subject matter, peer counselors exhibited language of attachment and involvement. These findings validate the research question of this study, and lead to other questions to answer. For

example, could these findings or differences be accounted for due to difference of style of the person, or is it a true reflection of the mind of the counselor? Could the manner of speech of the counselor been affected by the interviewer's educational level? Would these differences be found if testing counselors in other areas of the nation, instead of the border region? And lastly, would the significant categories vary if the sample would have been larger?

5.2 Conclusion

This study was able accomplish the purpose that it began with. The first one being: initiating a study involving the analysis of difference between peer and professional substance use disorder counselors. This was accomplished by bringing attention to the apparent disjunction of preference between peer and professional counselors in treatment facilities by addicts through the use of language. By researching studies that existed regarding efficacy and preference of one type of counselor over the other, this study was able to confirm that there still is not enough research on the matter and it is time that the research is invested on so it can be in a way resolved.

The second purpose this study accomplished was to investigate how the language of peer counselors in comparison to professional counselors was able to reveal a difference in state of mind of each. As mentioned in the introduction, the notion that language is the window to the mind of a person is not a new thought, it was first stated by Plato and in recent years has been talked about by psychologists and linguists alike. As Dr. Pennebaker states it "Words, in my world, are a window into the inner workings of people, a fascinating and revealing way to think about language and its links to the world around us all." (2011, p. 17). This study did just that. By using language analysis tools, this study was able to differentiate a total of 15 categories in

which peer and professional counselors differed significantly. These categories also showed that the distinction in the state of mind of these counselors were that of attachment and detachment of the situation. A more cognitive perspective on the matter at hand was used by the professional counselors, and a more emotional perspective was used by the peer counselors.

The last purpose that this study accomplished was to analyze the efficacy of language analysis tools to analyze natural speech without human intervention. This was particularly important regarding the use of LIWC, since, as this study discovered, the way the categories and the dictionaries for the categories were created were a bit flawed linguistically. It was proven that manual (human) intervention was needed in order to differentiate between different types of the same word that have different meanings, like the case of “you”. Without that intervention the results would have been incorrect and overgeneralized. Thus proving, that to this day, human intervention is still essential to the proper interpretation and analysis of language data, regardless of the amount of times or the size of dictionaries that a program may have.

Overall, although society has come a long way in terms of research and technology, there are still some areas where it is lacking. This study hopes that it will be able to serve as a jumping board to better the research concerning peer and professional substance use disorder counselors and their language differences and possible state of mind differences to aid the substance use disorder arena. It also hopes that the improving of language analysis tools continues so that one day linguistic or human intervention is unnecessary in language analysis, and that way language analysis can occur at faster rates to continue to learn about the “black box” of the mind.

5.3 Limitations of Study

Although this study accomplished many things, the areas for continuing study that were not answered and further analysis could explain, would be the possible sampling error that a study could have. Another would be that due to the size of our sample a better picture of the global group being discussed cannot be established. Further testing with a larger group sample could yield a more general statement as to whether the findings of this study can be attributed to all peer and professional counselors in other areas of the United States and or the world. Another limitation was that due to using secondary data analysis for this study, the data that was being analyzed was not collected specifically for the questions that this study asked. This researcher in a near future would like to tackle these limitations and see if the findings would change if these limitations would be taken into account.

Bibliography

AntConc 3.2 Multi Platform Corpus Analysis Toolkit

Barlow, M. (2003). *Concordancing and Corpus Analysis Using MP 2.2*. Houston: Athelstan.

Berdes, C. J. (2007). The Language of Caring: Nurse's Aides' Use of Family Metaphors Conveys Affective Care. *The Gerontologist*, Vol. 47, 340-349.

Bogenschutz, M. P., Geppert, C. M., & George, J. (2006). The role of twelve-step approaches in dual diagnosis treatment and recovery. *The American Journal on Addicction*, 15, 50-60.

Chung, C. K. (2008). Revealing dimensions of thinking in open-ended self-descriptions: An automated meaning extraction method for natural language. *Journal of Research in Personality* 42, 96-132.

Doukas, N., & Cullen, J. (2010, june). Recovered addicts working in the addiction field: Pitfalls to substace abuse relapse. *Drugs: education, prevention and policy*, 17(3), 216-231.

Hecksher, D. (2007). Former Substance Users Working as Counselors. A Dual Relationship. *Substance Use & Misus*, 42, 1253-1268.

Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2014). Pronoun Use Reflects Standings in Social Hierarchies. *Journal of Language and Social Psychology* , 125-143.

Kadushin, L., & Kadushin, A. (1969). The Ex-addict as a Member of the Therapeutic Team. *Community Mental Health Journal*, 5(5), 386-393.

Kasser, Tim. (2014, March). *Lucy in the mind of Lennon: An empirical analysis of Lucy in the Sky with Diamonds*. Retrieved from <http://thepsychreport.com/culture/lucy-in-the-mind-of-lennon-an-empirical-analysis-of-lucy-in-the-sky-with-diamonds/>

Krzywinski, M. (2012, Oct). *Lexical analysis of 2012 us presidential and vice-presidential debates*. Retrieved from <http://mkweb.bcgsc.ca/debates2012/>

Lakoff, G. a. (2003). *Metaphors We Live By*. Chicago: University of Chicago Press.

LIWC Inc. *What is LIWC?*

Lo Sciuto, L., Aiken, L., Ausetts, M., & Brown, B. (1984, May). Paraprofessional versus professional drug abuse counselors: attitudes and expectations of the counselors and their clients. *Int J Addict*, 19(3), 233-52.

McLellan, A. T., Luborsky, L., Woody, G., & Goebel, L. (1987). Counselor differences in methadone treatment. In L. S. Harris (Ed.), *Problems of Drug Dependence 1987: Proceedings of the 49th annual Scientific Meeting* (pp. 242-250). Rockville: NIDA Research Monograph 81.

McLellan, A. T., Woody, G. E., Metzger, D., McKay, J., Durell, J., Alterman, A. I., & O'Brien, C. P. (1996). Evaluating the Effectiveness of Addiction Treatments: Reasonable Expectations, Appropriate Comparisons. *The Milbank Quarterly*, 74(1), 51-85.

Mendoza, L. (2012). Stance and Patterns of Lexical Co-Occurrence in ESOL Students' Writings about Social Conflict and Violence. Extended Paper. University of Texas at El Paso.

- Najavits, L. M., & Weiss, R. D. (1994). Variations in therapist effectiveness in the treatment of patients with substance use disorder: an empirical review. *Addiction*(89), 679-688.
- Pennebaker, J. W. (2011). *The Secret Life of Pronouns: what our words say about us*. New York: Bloomsberry Press.
- Pennebaker, J. (2002). What Our Words Can Say About Us: Towards A Broader Language Psychology . *Psychological Science Agenda*.
- Snowden Jr, L., & Cotler, S. (1974). The effectiveness of paraprofessional ex-addict counselors in a methadone treatment program. *Psychotherapy: Theory, Research and Practice*, 11(4), 331-338.
- Tauszcisk, Y. A. (2010). The Psychological Meaning of Words: LIWC . *Journal of Language and Social Psychology and Computerized Text Analysis Methods*, 24-54.
- Wakeman, S. E. (2013, Feb 14). Letters: Language and Addiction: Choosing Words Wisely. *American Journal of Public Health*, pp. e1-e3.
- Weil, A. (2004). *The Natural Mind*. New York: Houghton Mifflin Company.
- Wexler, H. K., Falkin, G. P., & Lipton, D. S. (1990). Outcome evaluation of a prison therapeutic community for substance abuse treatment. *Criminal Justice and Behavior*, 17, 71-92.
- White, W. L. (2000). The history of recovered people as wounded healers: II. The era of professionalization and specialization. *Alcoholism Treatment Quarterly*, 18(2), 1-25.
- Ya Temkin, A. (2011). Consciousness, subconsciousness, theory of states of mind and its applications. *NeuroQuantology*, 9(4), 669-680. doi: ISSN 1303 5150
- Zuñiga, M. E. (1992). Using Metaphors in Therapy: Dichos and Latino Clients. *Social Work Vol.37, No. 1*, 55-60.

Appendix

Table 1 LIWC Dictionary sample

Table 1: LIWC2007 Output Variable Information

<i>Category</i>	<i>Abbrev</i>	<i>Examples</i>	<i>Words In Category</i>	<i>Validity (Judges)</i>	<i>Alpha: Binary/Raw</i>
Linguistic Processes					
Word count	wc				
words/sentence	wps				
Dictionary words	dic				
Words>6 letters	sixltr				
Total function words	funct		464		.97/.40
Total pronouns	pronoun	I, them, itself	116		.91/.38
Personal pronouns	ppron	I, them, her	70		.88/.20
1st pers singular	i	I, me, mine	12	.52	.62/.44
1st pers plural	we	We, us, our	12		.66/.47
2nd person	you	You, your, thou	20		.73/.34
3rd pers singular	shehe	She, her, him	17		.75/.52
3rd pers plural	they	They, their, they'd	10		.50/.36
Impersonal pronouns	ipron	It, it's, those	46		.78/.46
Articles	article	A, an, the	3		.14/.14
Common verbs	verb	Walk, went, see	383		.97/.42
Auxiliary verbs	auxverb	Am, will, have	144		.91/.23
Past tense	past	Went, ran, had	145	.79	.94/.75
Present tense	present	Is, does, hear	169		.91/.74
Future tense	future	Will, gonna	48		.75/.02
Adverbs	adverb	Very, really, quickly	69		.84/.48
Prepositions	prep	To, with, above	60		.88/.35
Conjunctions	conj	And, but, whereas	28		.70/.21
Negations	negate	No, not, never	57		.80/.28
Quantifiers	quant	Few, many, much	89		.88/.12

Numbers	number	Second, thousand	34		.87/.61
Swear words	swear	Damn, piss, fuck	53		.65/.48
Psychological Processes					
Social processes	social	Mate, talk, they, child	455		.97/.59
Family	family	Daughter, husband, aunt	64	.87	.81/.65
Friends	friend	Buddy, friend, neighbor	37	.70	.53/.12
Humans	human	Adult, baby, boy	61		.86/.26
Affective processes	affect	Happy, cried, abandon	915		.97/.36
Positive emotion	posemo	Love, nice, sweet	406	.41	.97/.40
Negative emotion	negemo	Hurt, ugly, nasty	499	.31	.97/.61
Anxiety	anx	Worried, fearful, nervous	91	.38	.89/.33
Anger	anger	Hate, kill, annoyed	184	.22	.92/.55
Sadness	sad	Crying, grief, sad	101	.07	.91/.45
Cognitive processes	cogmech	cause, know, ought	730		.97/.37
Insight	insight	think, know, consider	195		.94/.51
Causation	cause	because, effect, hence	108	.44	.88/.26
Discrepancy	discrep	should, would, could	76	.21	.80/.28
Tentative	tentat	maybe, perhaps, guess	155		.87/.13
Certainty	certain	always, never	83		.85/.29
Inhibition	inhib	block, constrain, stop	111		.91/.20
Inclusive	incl	And, with, include	18		.66/.32
Exclusive	excl	But, without, exclude	17		.67/.47
Perceptual processes	percept	Observing, heard, feeling	273		.96/.43
See	see	View, saw, seen	72		.90/.43
Hear	hear	Listen, hearing	51		.89/.37
Feel	feel	Feels, touch	75		.88/.26

Biological processes	bio	Eat, blood, pain	567	.53	.95/.53
Body	body	Cheek, hands, spit	180		.93/.45
Health	health	Clinic, flu, pill	236		.85/.38
Sexual	sexual	Horny, love, incest	96		.69/.34
Ingestion	ingest	Dish, eat, pizza	111		.86/.68
Relativity	relativ	Area, bend, exit, stop	638		.98/.51
Motion	motion	Arrive, car, go	168		.96/.41
Space	space	Down, in, thin	220		.96/.44
Time	time	End, until, season	239		.94/.58
Personal Concerns					
Work	work	Job, majors, xerox	327		.91/.69
Achievement	achieve	Earn, hero, win	186		.93/.37
Leisure	leisure	Cook, chat, movie	229		.88/.50
Home	home	Apartment, kitchen, family	93		.81/.57
Money	money	Audit, cash, owe	173		.90/.53
Religion	relig	Altar, church, mosque	159		.91/.53
Death	death	Bury, coffin, kill	62		.86/.40
Spoken categories					
Assent	assent	Agree, OK, yes	30		.59/.41
Nonfluencies	nonflu	Er, hm, umm	8		.28/.23
Fillers	filler	Blah, lmean, youknow	9		.63/.18

Words in category refers to the number of different dictionary words that make up the variable category. **Validity judges** reflect the simple correlations between judges' ratings of the category with the LIWC variable (from Pennebaker & Francis, 1996). **Alphas** refer to the Cronbach alphas for the internal reliability of the specific words within each category. The binary alphas are computed on the occurrence/non-occurrence of each dictionary word whereas the raw or uncorrected alphas are based on the percentage of use of each of the category words within the texts. All alphas were computed on a sample of 2800 randomly selected text files from our language corpus.

The LIWC dictionary generally arranges categories hierarchically. For example, all pronouns are included in the overarching category of function words. The category of pronouns is the sum of personal and impersonal pronouns. There are some exceptions to the hierarchy rules:

Common verbs are not included in the function word category. Similarly, common verbs (as opposed to auxiliary verbs) that are tagged by verb tense are included in the past, present, and future tense categories but not in the overall function word categories.

Social processes include a large group of words (originally used in LIWC2001) that denote social processes, including all non-first-person-singular personal pronouns as well as verbs that suggest human interaction (talking, sharing).

Perceptual processes include the entire dictionary of the Qualia category (which is a separate dictionary), which includes multiple sensory and perceptual dimensions associated with the five senses.

Table 2 *ProvP-Ant*

Hits		Total No. of Word Types: 3002		Total No. of Word Tokens: 34835
Rank	Freq	Word	Lemma Word Form(s)	
1	1218	the		
2	1205	and		
3	1124	to		
4	1032	i		
5	970	that		
6	809	we		
7	761	of		
8	746	you		
9	735	a		
10	666	they		
11	579	it		
12	486	in		
13	485	is		
14	474	have		
15	456	are		
16	398	s		
17	358	know		
18	332	so		
19	296	t		
20	286	for		
21	281	them		
22	270	do		
23	257	what		
24	246	but		
< >	< >	< >	<	

Table 3 *Pro-Ant*

Hits		Total No. of Word Types: 1959		Total No. of Word Tokens: 17919
Rank	Freq	Word	Lemma Word Form(s)	
1	682	the		
2	604	to		
3	556	that		
4	548	and		
5	456	we		
6	454	i		
7	423	of		
8	414	they		
9	387	a		
10	302	it		
11	296	have		
12	293	you		
13	233	in		
14	231	is		
15	228	are		
16	187	so		
17	177	s		
18	174	know		
19	171	them		
20	158	if		
21	152	t		
22	151	do		
23	142	or		
24	138	for		
< >	< >	< >	< >	

Table 4 *P-ant*

Hits		Total No. of Word Types: 2013		Total No. of Word Tokens: 16916
Rank	Freq	Word	Lemma Word Form(s)	
1	657	and		
2	578	i		
3	536	the		
4	520	to		
5	453	you		
6	414	that		
7	353	we		
8	348	a		
9	338	of		
10	277	it		
11	254	is		
12	253	in		
13	252	they		
14	228	are		
15	221	s		
16	184	know		
17	178	have		
18	156	what		
19	148	for		
20	145	so		
21	144	t		
22	124	with		
23	119	do		
24	119	was		
< >	< >	< >	< >	

Table 5 LIWC Pro analysis

Filename	Segment	funct	pronoun	ppron	i	we	you	shehe	they	ipron	article	verb
T1	1	63.36	19.03	12.2	4.35	2.67	1.29	0.27	3.61	6.83	6.83	17.18
T3	1	65.12	20.05	12.63	4.75	0.52	1.56	0.29	5.5	7.42	8.29	15.47
T4	1	62.55	18.98	12.99	2.46	4.35	1.95	0.13	4.1	5.99	5.36	15.26
T5a	1	60.04	19.27	10.75	2.02	3.36	1.64	0.07	3.66	8.51	5.9	17.18
T6	1	60.26	16.98	8.58	2.95	1.95	1.04	0.91	1.73	8.4	7.22	14.62
T11	1	64.43	21.9	14.64	1.9	3.46	4.04	0.24	5	7.26	5	18.77
T12	1	59.4	18.88	11.51	1.62	4.7	0.32	0.24	4.62	7.37	7.46	17.83
T13	1	63.85	21.13	13.43	2.63	4.98	0.94	0.47	4.41	7.7	5.82	18.22
T14	1	62.9	19.08	11.01	1.35	4.35	0.97	0.1	4.25	8.07	5.7	16.04
auxverb	past	present	future	adverb	preps	conj	negate	quant	number	swear	social	family
10.79	1.61	12.67	1.37	4.16	12.24	7.77	3.06	4.24	0.51	0	13.18	0.63
9.1	0.81	13.09	0.41	4	12.98	8.23	2.03	2.38	1.51	0	13.33	0.75
8.95	0.63	11.66	1.32	6.62	13.56	9.65	1.01	2.33	0.82	0	17.28	0.88
9.71	0.67	14.56	0.82	4.78	10.68	9.63	1.42	1.87	0.75	0	15.53	1.05
8.95	1.23	11.35	0.95	4.04	13.99	6.86	1.41	3.41	0.41	0	11.67	0.82
11.05	1.87	14.76	1.08	5.3	11.69	8.92	2.23	1.9	0.63	0.03	18.46	0.81
9.97	1.78	13.61	1.3	3.32	11.26	8.43	0.97	1.86	0.73	0	14.91	0.65
11.83	1.22	12.68	1.88	5.45	11.55	9.11	1.97	2.44	0.38	0	14.37	0.19
9.32	1.64	12.27	1.26	5.75	12.17	10.1	1.3	2.75	0.34	0	14.69	1.01
friend	humans	affect	posemo	negemo	anx	anger	sad	cogmech	insight	cause	discrep	tentat
0	1.45	2.71	1.96	0.71	0.04	0.24	0.2	20.99	3.69	1.96	2.04	4
0.06	1.27	3.24	1.91	1.04	0.41	0.17	0.17	20.28	3.42	2.9	1.8	3.71
0.06	0.95	2.59	1.83	0.69	0.13	0.06	0.06	21.75	3.15	1.39	2.52	4.35
0.15	1.05	4.18	3.44	0.52	0	0	0.15	22.55	4.03	2.17	2.69	3.21
0	2.45	4.13	1.91	2.23	0.77	0.68	0.05	18.35	3.22	1.5	1.41	3.95
0	0.81	3.04	2.23	0.72	0.15	0.27	0.18	22.08	4.01	2.23	2.47	4.31
0	0.97	4.21	3	1.05	0.16	0.16	0.24	21.39	2.67	2.67	3.24	2.76
0	1.31	2.63	1.69	0.94	0	0.47	0.28	20.19	3	1.13	2.35	3.29
0	0.68	3.82	2.9	0.92	0.1	0.29	0.14	23.77	2.27	2.61	2.75	5.27
certain	inhib	incl	excl	percept	see	hear	feel	bio	body	health	sexual	ingest
1.1	0.39	6.08	4.59	1.73	0.78	0.86	0.04	0.78	0	0.63	0.08	0.12
0.81	0.7	5.85	3.65	1.39	0.52	0.52	0.23	1.62	0.17	0.64	0.06	0.93
0.76	0.57	8.39	4.1	0.82	0.13	0.44	0.19	1.45	0	0.63	0	1.01
1.57	0.22	8.66	2.99	1.19	0.67	0.3	0.22	1.12	0.15	0.97	0	0
0.82	0.14	6.31	2.86	1.23	0.68	0.36	0.14	2.09	0.18	1.63	0.14	0.82
0.78	0.54	7.14	4.16	1.3	0.39	0.6	0.27	1.08	0.06	0.66	0.06	0.39
0.65	0.49	8.67	3.24	1.46	0.73	0.16	0.57	1.05	0	0.73	0	0.49
1.22	0.28	7.98	2.91	0.47	0.19	0.28	0	1.22	0.09	0.94	0.09	0.19
1.4	0.58	7.83	5.31	2.03	0.92	0.77	0.34	1.3	0	0.92	0.14	0.58

relativ	motion	space	time	work	achieve	leisure	home	money	relig	death	assent	nonfl
9.49	2.08	4.08	3.84	1.92	1.88	0.43	0.35	0.24	1.26	0.08	0.43	0.75
10.66	2.32	3.53	4.81	1.68	2.55	1.45	0.75	0.06	0.98	0	0.46	0.41
11.92	1.83	5.55	4.48	2.46	1.39	0.69	1.13	0.19	1.2	0	0.38	0.69
8.51	1.87	3.36	3.14	4.63	3.06	0.97	1.27	0.15	1.19	0.07	0.37	0.52
10.45	1.95	5.86	2.18	2.36	1.59	1.14	0.91	0.14	0.18	0	0.59	1.54
10.51	2.23	4.55	3.49	1.99	1.33	1.23	0.81	0.21	0.78	0.06	0.66	0.39
10.37	1.78	4.86	2.76	2.51	1.86	1.05	0.89	0.49	1.13	0	0.73	0.08
13.33	2.25	5.73	5.07	0.94	0.85	0.56	0.66	0.56	0.09	0.38	0.38	0.85
8.94	1.79	4.15	2.8	3.82	1.4	1.3	1.16	0.14	0.43	0	0.68	0.48

filler
0.55
0.17
0.19
0.22
0.09
0.15
0
0.19
0.14

Table 6 LIWC Peer analysis

Filename	Segment	funct	pronoun	ppron	i	we	you	shehe	they	ipron	article	verb
T5b	1	62.88	20.98	11.88	1.71	1.6	4.61	0.41	3.55	9.1	5.2	19.39
T5c	1	59.83	20.52	14.41	4.37	1.75	5.24	0	3.06	6.11	4.37	20.09
T7	1	63.61	20.71	14.57	5.04	2.82	1.95	0.64	4.13	6.14	5.61	16.08
T8	1	60.66	19.8	11.36	4.58	1.26	1.99	0.27	3.26	8.44	5.38	17.81
T9	1	59.61	18.29	11.16	2.71	4.2	1.66	0.78	1.81	7.13	5.74	14.58
T10a	1	64.53	19.02	11.46	1.83	2.53	4.65	1	1.45	7.56	6.27	16.99
T10b	1	61.63	22.18	15.97	8.2	1.2	3.51	1.36	1.7	6.2	4.93	17.33

auxverb	past	present	future	adverb	preps	conj	negate	quant	number	swear	social	family
13.48	0.77	16.25	1.12	4.96	9.63	7.15	2.19	2.3	0.47	0	15.13	0.24
13.1	2.18	15.28	0.87	4.37	9.61	6.99	3.49	3.06	0	0	13.54	0.44
9.23	3.96	9.77	0.44	5.4	12.35	9.33	1.48	2.22	0.91	0	16.58	0.67
10.43	0.33	14.68	1.4	4.39	10.17	8.31	3.79	3.19	0.2	0.07	12.69	0.66
8.55	1.95	10.26	0.76	4.22	12.58	9.18	1.27	2.32	0.78	0	14.43	0.39
10.59	1.83	13.75	0.25	5.77	13	8.31	2.16	2.24	1.08	0.04	14.95	0.66
9.24	4.71	10.76	0.4	5.38	11.74	7.69	1.3	2.1	0.8	0.05	14.4	1.2
friend	humans	affect	posemo	negemo	anx	anger	sad	cogmech	insight	cause	discrep	tentat
0	2.25	3.25	1.83	1.42	0.65	0.06	0.35	18.74	2.48	3.43	2.19	3.01
0	1.31	5.68	3.93	1.75	0.87	0.44	0.44	17.9	0.87	3.06	0.44	0.44
0	2.11	2.85	2.25	0.57	0.27	0.03	0.07	19.3	2.52	2.08	1.04	2.42
0.07	1.59	3.19	2.19	1	0.53	0.07	0	22.72	4.05	1.99	3.12	6.11
0.1	1.27	2.81	1.88	0.93	0.15	0.32	0.15	22.34	3.15	1.25	1.37	2.95
0	1.99	2.41	1.41	1	0.08	0.37	0.29	19.19	2.33	2.45	1.41	2.41
0.19	2.08	3.7	1.65	2.05	0.29	0.93	0.27	16.56	2.8	1.94	1.17	1.25
certain	inhib	incl	excl	percept	see	hear	feel	bio	body	health	sexual	ingest
0.59	0.71	5.08	3.25	1.95	0.53	0.95	0.47	0.95	0.06	0.71	0.12	0.18
1.75	1.31	6.99	3.49	2.62	0.87	1.31	0	1.75	0	1.31	0	0.44
1.31	0.07	8.22	2.72	1.58	0.5	0.67	0.4	1.75	0.13	0.97	0.1	0.64
1.13	0.73	4.92	4.25	2.06	0.8	0.66	0.6	1.79	0.2	0.93	0.13	0.86
1.1	0.29	10.09	2.76	1.27	0.32	0.78	0.12	1.76	0.2	1.27	0.1	0.61
0.66	0.25	8.1	2.82	1.79	0.29	1.29	0.12	1.54	0.42	0.79	0.21	0.21
1.04	0.32	7.03	1.76	1.36	0.61	0.64	0.11	2.24	0.19	1.62	0.08	0.53
relativ	motion	space	time	work	achieve	leisure	home	money	relig	death	assent	nonfl
9.75	2.54	3.37	3.78	0.83	1.3	0.3	0.3	0.3	0.71	0.06	0.65	0.59
11.35	3.49	3.93	4.8	0	1.31	1.75	0.87	0.44	0	0	3.06	1.75
12.69	1.91	4.33	6.08	2.15	1.48	1.17	1.14	0.07	2.22	0.03	0.27	1.07
8.77	1.13	3.72	4.32	1.59	1.46	1.13	0.66	0.2	0.86	0	1.13	0.27
11.09	1.39	6.4	3.22	2.15	3.27	0.83	0.71	0.32	0.24	0.2	0.59	2.49
12.5	2.95	5.32	3.78	1.7	1.45	0.62	0.79	0.08	1.25	0	0.17	0.08
11.42	2.4	4.9	3.89	1.22	1.28	0.75	0.43	0.27	1.76	0.19	0.45	0.19

filler
0.24
0.44
0.2
0.66
0.29
0.21
0.11

Table 14 Reliability Checking Errors

Peer Interviews Transcript Errors		Professional Transcript Errors	
Interview #	Error #	Interview #	Error #
T5b	6	T1	8
T5c	0	T3	18
T7	7	T4	3
T8	1	T5a	2
T9	2	T11	4
T10a, T10b	10	T12	1
		T13	4
		T14	8
Total Errors		TOTAL:	48
	26		

Vita

Lizzeth C. Jensen, raised in Miami, Florida, graduated with honors from Miami South Ridge Senior High school. After graduation she pursued her Bachelors of Arts at the University of Florida. Due to family issues, she moved to Texas her senior year at the University of Florida. Once in El Paso, she graduated from The University of Texas at El Paso with Honors with a Bachelors of Arts in Linguistics. Upon graduation she was awarded the Anna J. Cooper Award for Outstanding Academic Achievement. A few months after graduation, she applied and was accepted to the Graduate Linguistics program at The University of Texas at El Paso. During her time there she was awarded a Research fellowship from the Vulnerability Issues in Drug Abuse (VIDA) program under the mentoring of Dr. Mary Cuadrado. She was also a part of the Language Acquisition and Linguistics Research Lab. During the time of her graduate career, Lizzeth was also a Spanish instructor for Berlitz, Inc.

Permanent address: 9146 Texas Red Dr
El Paso, TX. 79907

This thesis was typed by Lizzeth C. Jensen