Linguistic Features of False Confessions and Confessions Not in Dispute: A Corpus Analysis

Lucrezia Rizzelli
CUNY John Jay College, lucrezia.rizzelli@gmail.com

Follow this and additional works at: https://academicworks.cuny.edu/jj_etds
Part of the Criminology and Criminal Justice Commons, Law and Psychology Commons, Other Psychology Commons, and the Psycholinguistics and Neurolinguistics Commons

Recommended Citation
https://academicworks.cuny.edu/jj_etds/117

This Thesis is brought to you for free and open access by the John Jay College of Criminal Justice at CUNY Academic Works. It has been accepted for inclusion in Student Theses by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.
Linguistic Features of False Confessions and Confessions Not in Dispute: A Corpus Analysis

Lucrezia Rizzelli

John Jay College of Criminal Justice
This thesis is dedicated to my family. To my grandparents, Anna and Michele, for always seeing the best in me (nonno, “colpo da maettro!”). To my mom, for teaching me to dream big and for supporting me even when my dreams brought me one ocean away; and to my dad, for being my hero and for never letting me peek into the Latin dictionary when a complete translation was there. I am but a dwarf on the shoulders of a giant.
Table of Contents

Introduction ......................................................................................................................... 5

Method ............................................................................................................................... 11
  Sample .............................................................................................................................. 11
  Materials .......................................................................................................................... 12
  Procedure ......................................................................................................................... 14

Results .................................................................................................................................. 16
  Baseline Characteristics of the Corpora ........................................................................... 16
  Differences Between the Corpora .................................................................................... 18
  Predictors: Linguistic Function and Context .................................................................... 19

Discussion .......................................................................................................................... 21
  Limitations and Future Research ...................................................................................... 25

References .......................................................................................................................... 28

Tables and Figures .............................................................................................................. 34

Appendix A ......................................................................................................................... 40
Abstract

Confessions are considered the gold standard of evidence, and yet many cases of false confessions causing wrongful convictions have come to the surface in the past decades. Currently, a method to identify false confessions does not exist and studies focusing on the content of the confessions have found similarities rather than points of distinction. In this study, we approached confessions from a stylistic rather than qualitative point of view, utilizing corpus analysis to outline the linguistic features of two samples of confessions: false confessions \( (n=37) \) and confessions not in dispute \( (n=98) \). Subsequently, we created a model through logistic regression in order to distinguish the two, based on three predictors: personal pronouns, impersonal pronouns and conjunctions. In a first sample comprised of 25 false confessions and 25 confessions not in dispute the model correctly categorized 37 out of 50 confessions \( (74\% \) accuracy rate), and in a second out-of-model sample the predictors accurately classified 20 of the 24 confessions \( (83.3\% \) accuracy rate). A high frequency of impersonal pronouns was found to be associated with a higher likelihood of the confession being false, while confessions containing a higher rate of conjunctions and personal pronouns were more likely to belong to the sample not in dispute. Lastly, in confessions not in dispute “I” was often followed by a lexical verb, a pattern that was found with less frequency in false confessions. Disparities were also found in the way variations of the sentence “I don’t remember” were used within the corpora.

Keywords: Confessions, Corpus Linguistics, Interrogations, False Confessions, Linguistics
The existence of false confessions has for decades remained concealed by the mistaken “common sense” assumption that innocents simply do not confess to crimes they did not commit (Kassin, 2012). Yet, out of the 364 individuals incarcerated for murder and rape that the Innocence Project has helped exonerate through postconviction DNA testing, 28% of the cases contained a false confession as a contributing factor; and many confessions were also “corroborated” by equally erroneous statements made by their co-defendants (https://www.innocenceproject.org). A lower estimate is offered by the National Registry of Exoneration, a broader and more diverse sample of wrongful convictions, which reports that out of the approximately 2,400 exonerees in their database, 291 had falsely confessed, amounting to 15% of all cases (National Registry of Exoneration, 2019).

Confessions are considered the gold standard when it comes to evidence (Kassin, 1997; 2012). In *Bruton v United States* (1968), the Supreme Court claimed that confession evidence is “probably the most probative and damaging” (p.7). Their power in court is supported by the weight they add toward the final verdict of guilt to the point that, as experiments have presented, confession evidence influences jurors’ decision more heavily than eyewitness identification and character evidence (Kassin & Neumann, 1997). In fact, research has shown that if a false confessor rejects a guilty plea and opts for trial, the chances of conviction range between 73% and 81% (Leo & Ofshe, 1998; Drizin & Leo, 2004). Moreover, a study in which jurors were presented with evidence of coercion in the interrogation process, hence discounting the confession, showed that 44% of them were still swayed by it toward a judgment of guilt, compared to the 19% of the control group (Kassin & Sukel, 1997). Indeed, a similar pattern was found in a study of judges (Wallace & Kassin, 2012). A crucial aspect regarding confessions is that people are unable to distinguish between true and false confessions, independently by their level of expertise and experience in the field. In a study by Kassin, Meissner and Norwick (2005) police investigators and college students watched or listened to
ten inmates’ confessions (half were true, and half were false) and college students discriminated the two with more accuracy (53.4% for videotapes and 64.1% for audiotapes) than police investigators (42.1% for videotapes and 54.5% for audiotapes). The investigators also showed more confidence in their judgements and a bias towards guilt presumption. When in a second study the participants were informed that half the confessions were true and half were false, the guilt presumptive bias among investigators was eliminated, although it did not improve the accuracy or decrease their confidence. Similar results were replicated in a study focused on the ability of laypersons to evaluate the credibility of confessions given by juveniles, which averaged a 52.8% accuracy rate among college students, with increased rates in case of an audio or video recording over a transcript (Honts, Kassin, & Craig, 2014).

While a confession is defined as “an admission of guilt followed by a narrative statement of what, how, and why the confessor committed the crime” (Kassin, Perillo, Appleby, & Kukucka, 2015), false confessions have been divided in three categories: voluntary (not induced by the interrogation), coerced compliant (when people confess during the interrogation to reduce the stress of it, to avoid any sort of punishment, or in exchange of a promised reward), and coerced-internalized (when an individual comes to the belief that they have indeed committed the crime they are accused of, which can also be erroneously be strengthened by false memories) (Kassin & Wrightsman, 1985). Since research began shedding light on how innocent people can be induced to confess, the factors contributing to this phenomenon started to come to the surface, to the point that in 2008 the American Psychological Association (APA) released an amicus curiae brief in reaction to Wright v. Commonwealth of Pennsylvania, stating that it is possible for a suspect to produce a confession in response to coercive police interrogations techniques (Kassin, 2012). Shortly thereafter, in 2010, the American Psychology-Law Society published a White Paper, only the second in its history, titled “Police-Induced Confessions: Risk Factors and Recommendations” (Kassin,
Drizin, Grisso, Gudjonsson, Leo, & Redlich, 2010). Therefore, when a court needs to decide whether a confession was voluntary, and hence admissible as evidence, or coerced, and hence not admissible, two sets of factors are analyzed: (1) the totality of the relevant circumstances regarding the suspect and interrogation, and (2) the statement made by the defendant, to determine whether it is reliable (Culombe v. Connecticut, 1961). In order for the trial judge to determine the voluntariness of the confession, they shall consider age and mental state of the suspect, whether the suspect was in custody and apprised of his or her rights, the time elapsing between the arrest and the confession, and the tactics police used in their interrogation (18 U.S. Code § 3501). Still, some psychological aspects of the interrogation process exerting an influence on the suspect (e.g. feedback style, desire to please), especially in case of a juvenile or of a person affected by a mental disorder, are not considered within the spectrum of “circumstances” (Dassey v. Dittman, 2017).

Past studies have concentrated their scope mainly on the qualitative contents of the confessions. Garrett (2010) examined 38 false confessions and pointed out how the interrogation process could contaminate confession statements, with the inclusion of facts that where kept out of the public domain and that supposedly only the person who had committed the crime could have been privy to (found in 36 of the 38 confessions). In a second analysis, Garrett (2015) analyzed 66 additional confessions and found a similar pattern of statements contaminated with inside information in 94% of the cases. Contamination during the interrogation process is as risky as the one concerning physical evidences of the crime scene, as it still represents an “unwanted transfer of material” from a source to another (National Institute of Justice, 2000). In the case of false confessions, contamination occurs mostly during the interrogation, when police purposefully or inadvertently communicate nonpublic information regarding the crime to the suspect, therefore rendering the final confession the product of the police’s leading questions or suggestions, even despite the training police
officers receive in order to withhold nonpublic details from suspects during the interrogations (Nirider, Tepfer, & Drizin, 2012).

Appleby, Hasel, and Kassin (2013) conducted a content analysis on a sample of 20 false confessions and found out that they all contained information regarding the defendant’s purported thoughts and feelings at the time of the offense, as well as factual details. They all reported time and location of the crime, visual details, and references to the victim and the victim’s behavior; additionally, almost all the confessions contained references to auditory sensations, to other people, and the to victim’s appearance/mental state. Moreover, this study underlined the presence of apologies, minimizations themes, assertions of voluntariness, and remorse.

The richness and detailed nature of these confessions are part of the reason why both juries and judges weigh confession evidence heavily in their conviction (Appleby et al., 2013; Drizin & Leo, 2004; Kassin & Sukel, 1997; Wallace & Kassin, 2012). In order to better understand how to identify questionable confessions, it would help to know what the typical “baseline” confession that is not in dispute sounds like. In a recent study, Appleby and colleagues (unpublished) conducted a content analysis on confessions that were not in dispute, which were subsequently compared to a corpus of 20 proven police-induced false confessions. The results showed a marked similarity between the two samples in regard to amount of detail and indicators of credibility, such as expressions of voluntariness, error corrections, apologies, and expressions of remorse.

Along parallel lines, the present study aims to analyze the language of confessions using an approach called corpus linguistics, which in contrast to content analysis examines how things are said rather than what is said (Baker, 2006). Exploring the linguistic patterns of the confessors’ speeches offers the possibility to detect distinguishing factors without having to rely on the content of the confessions, which can be influenced both by the suspect’s style of
speaking and police during the interrogation process. One advantage of using corpus analysis is that it is objective and quantitative, and hence less susceptible to the possible bias associated with the more subjective codings of content (Baker 2006). A second advantage is the possibility to discover systematic patterns in a large number of texts, which allows the detection of underlying discourses shared in a specific community (Stubbs, 2001). In linguistics, the word corpus identifies a collection of texts, either in the written or spoken form, usually stored on a computer (Romero-Trillo, 2013), which is a representative sample of a specific kind of language which occurs naturally in a certain context, hence why they are used as reference to examine the language of interest (Baker, 2006). Another important aspect of corpus analysis is that it is performed on the entire corpus rather than on the isolated items which compose it (Baker, 2006).

Corpus analysis has been used for a variety of purposes, from the creation of dictionaries (Clear, Fox, Francis, Krishnamurthy & Moon, 1996) to the study of how language varies over the years (Biber, 1991). In other cases, it has proved efficient at discriminating old versus young participants based on the variety of their lexicon (Horton, Spieler, & Shriber, 2010). It has also been used to examine the linguistic patterns of many corpora, from application letters (Henry & Roseberry, 2001) to medical language (Wu et al., 2012), to even music lyrics (De Clercq & Temperley, 2011).

In linguistics, the most common type of corpus analysis entails the comparison of the corpus of study against a reference corpus, like the Corpus of Contemporary American English or the Oxford English Corpus (Romero-Trillo, 2013). In the present study, there will not be a reference corpus, but rather a first delineation of baselines for the two corpora (false confessions and confessions not in dispute) followed by a comparison of the corpora against each other. For this type of analysis, a strong homogeneity between the corpora is a prerequisite to avoid the influence of unwelcomed third variables (Granger & Leech, 2014), such as
different speech formats or the theme and aim of the discourse. Thus, since the Innocence Project works on rape and murder cases, we only gathered confessions of such crimes from the FBI database, to avoid any modulating effect the typology of crime may cause.

Researchers have analyzed the language of confessions through different methods. In his book, *The Language of Confessions, Interrogation, and Deception*, (1998) Roger Shuy investigated the validity of confession evidence and the presence of possible coercive methods in specific cases, portraying his work as an expert witness. Lowrey and Ray (2015) analyzed six confessions (three false, three they labeled as true) using narrative analysis and found out that statements within the true confessions sample showed narratives minimizing blame and focusing on the events of the crime. In contrast, false confessions were characterized by more explanations of the motive without an evaluation of the crime itself. Another study showed that false confessions presented fewer adjectives than confessions defined as true, but no differences were found for verbs as indicators of deception (Villar, Arciuli, & Paterson, 2013). Lastly, denials of having committed the crime contained fewer modal verbs, fewer past tense verbs, and more present tense verbs than confessions (Ali & Levine, 2008).

We conducted a study with two main goals in mind. First, we aimed to define a linguistic baseline of confessions that are not in dispute as to their credibility by analyzing hundreds of law enforcement case files from a national sample. Secondly, we sought to compare the results to a smaller sample of proven false confessions.

This study was exploratory in nature, designed to describe the linguistic baselines of false confessions and confessions not yet in dispute from a stylistic point of view. Currently, there is no instrument or test able to discern these two types of confessions. Since it is now clear how much the final verdict and sentencing are influenced by confession evidence, it is of paramount importance to describe their defining characteristics and examine whether they differ stylistically.
Method

Sample

We gathered a sample of false confessions (FC) with assistance from the Innocence Project, which also provided a list of 82 DNA exonerees who had falsely confessed. It was not possible to retrieve the written confessions for each and every case within this Innocence Project sample, as confession evidence was not always present in their case files. The confessors from the Innocence Project had all confessed to rape and/or murder and had spent an average of 14 years in prison before being exonerated (https://www.innocencerecord.org). Additional FC were gathered from Dr. Brandon Garrett’s online database (DNA Exoneration Database, 2019) from files Dr. Saul Kassin’s (Personal Communication) had acquired through research and consulting on cases involving wrongfully convicted individuals who had confessed during his years of researching this phenomenon, and from Dr. Tammy Gales’ (Personal Communication) previous studies in the field. In total, the FC sample consisted of 40 false inculpatory statements. We divided these confessions in subsamples according to their format: Q&A (27), first-person statements (10), and trial proceedings (3). This last subsample was excluded from the analyses as this format was not present within the control group of confessions not in dispute and the confession evidence was at times presented as summary from the officer and therefore was not representative of the suspect’s speech.

We gathered a sample of confessions that were not contested and are not in dispute (CNID) from an FBI database housed on the ninth floor at John Jay College of Criminal Justice. The data for this research were taken from closed, fully adjudicated state and local cases that were contributed from law enforcement agencies from around the country for the purpose of research. All identifiers, including names of victims, suspects, offenders, officers, departments, and correctional agencies, are removed. Only aggregated data are reported on.

Even though the database features a broad variety of crimes, we only took into consideration cases falling in the spectrum of homicide and/or rape, as these were the crimes
present in the FC sample. The confessions from this national sample belong to the following criminal categories: single victim homicide, serial homicide, multiple homicide, single rape, serial rape, domestic homicide, and serial sexual homicide. This total sample consisted of 98 confessions, which were also divided into the two subgroups: Q&A \((n = 32)\) and first-person statements \((n = 66)\). Due to time constrains, we included only interview transcripts or statements that were less than 40 pages long. Also excluded were third-person statements, secondhand summaries of confessions within police reports, and confessions given in a language other than English.

**Materials**

The first software we used is a psychological language analysis program called Linguistic Inquiry and Work Count (LIWC)(Pennebaker, Booth, & Francis, 2015), and it was created to identify emotional, cognitive and structural components of language, both spoken and written. Initially, its main application was the study of language and disclosure (Francis, 1993; Pennebaker, 1993), but through the years and with the development of updated versions, it has found multiple applications. For instance, it has been used to compare use of personal pronouns in positive and negative political ads (Gunsch, Brownlow, Haynes, & Mabe, 2000), and to assess emotion and the degree of immersion victims of intimate partner violence showed during the narration (Holmes et al., 2007). LIWC operates by comparing the words in the corpus (called target words) to a list of words that are part of its internal dictionary (composed of approximately 6,400 words). Every target word present in the dictionary is sorted in one of the 95 categories representing different linguistic constructs (e.g. pronouns, conjunctions, punctuation), and the final output summarizes which percentage of the overall corpus falls under a specific category. As LIWC does not carry out analyses for statistical significance, we exported its output into SPSS, where we first conducted bivariate correlational analyses to draw a baseline of what FC and CNID sound like. Afterwards we used bivariate correlation to screen
the 95 variables into a subset of variables that had a higher likelihood of being predictive of either FC or CNID. Once these predictors were identified, we used logistic regression to determine whether a model could be generated which would be able to distinguish between FC and CNID. We chose LIWC because it is a psychological language analysis software which links linguistics to psychological constructs such as thought processes, emotions, motivations, and personality (Pennebaker, Boyd, Jordan, & Blackburn, 2015). Once we identified which of the 95 LIWC categories were predictors of either FC or CNID, we looked into the frequency distribution of the words within each of these categories in the corpora to identify which words had appeared the most in each corpus.

In order to conduct a qualitative analysis of the data and contextualize the results, we used a second software, a freeware corpus analysis program called AntConc (Anthony, 2014). The name of the software derives from Anthony, the last name of its creator, and concordancing, an operation through which the software identifies specific words within the entirety of the corpus, and where they appear in each text. We chose AntConc because it reports how many unique words appear in each corpus, which is a proxy for the level of lexical variety and vocabular complexity, and because the concordance tool allows in-depth analysis of the context and function of the words since it lists not only every instance in which a specific word appeared, but also the sentence and the larger context in which it is present (Anthony, 2014). We typed into the concordance search bar the word which was found to be most frequent for the predictive category, then used the sort settings to alphabetically list every instance starting from the first word positioned on the right of the node word in the search bar. Alphabetic sorting was also applied to the second and third position on the right from the lemma of interest. Once the output appeared, we scrolled through all the instances and annotated the raw frequencies of patterns that appeared with consistency. For example, when looking into the collocates for “I”, we would count how many times the sentence “I remember”, or “I killed her/him” appeared.
Other examples could involve modal verbs and their use, like “I should” or “I will”. To make sure all the hits were not coming from a single or few confessions, thus making the pattern subjective rather than systematic, we also used the concordance plot tool, which reports in how many texts within the corpus such a pattern is present. These patterns of words were typed into a spreadsheet and clustered in grammatical categories, such as modal verbs, lexical verbs, cognitive verbs etc. depending on the words that followed “I”. As the goal was to compare the usage of the word between the two corpora, the raw counts were standardized into percentages by dividing them for the overall words in the corpus.

Since the concordance tool lists the instances without offering descriptive statistics or the possibility to analyze the data, they were exported and further analyzed through SPSS. Researchers have used AntConc in many different linguistic contexts. For example, to analyze the writing styles of climate change proponents and skeptics (Medimorec & Pennycook, 2015), to explore how transgender individuals are represented in the British Press (Baker, 2014), and to examine the differences in the registers of official European Union texts and online news (Jablonkai, 2009).

**Procedure**

We gathered the FCs from the Innocence Records. The cases were screened based on the information provided by the abstracts on the website, which state whether or not a confession is present in the system. Since AntConc necessitates files in .txt format, we typed the PDFs of the FC into .doc documents. During this process, we kept typos and grammatical mistakes untouched because they represent specific aspects of speech, but we included the correct diction in brackets, so that their presence would be detected by the software when inputting the correct spelling.

In regard to the CNID, we performed different actions according to the style and format. We manually typed handwritten or short typed confessions into a computer in .doc format,
while we scanned long typed confessions and ran them through an optical character recognition (OCR) tool already present in Adobe Reader, which transforms PDF files into readable documents. We hand-checked all OCR texts for accuracy and used the same criterion concerning typos, as they were left untouched, with the addition of the correct spelling in brackets.

Subsequently, we split confessions in Q&A format from both samples between interrogator and suspect’s speech, in order to be able to isolate the speakers. Both sets of .doc files were then translated into .txt in order to be logged into AntConc.

To establish a baseline for linguistic characteristics of FC and CNID, we first used AntConc was to determine the ratio of unique words over total words in each corpus, to understand the level of complexity of the texts, as the ratio of unique words represents how broad the vocabulary of the speaker is. Then, we ran both corpora through LIWC, converted the percentages in LIWC’s output into raw counts, and used bivariate correlation analysis to examine which of the 95 categories provided by the software were associated with FC and which with CNID. We considered as significantly correlated only categories that showed a $p < 0.05$ and a $r > 0.2$.

To compare the types of confessions, we created a random sample of 25 FC and 25 CNID, controlling for format by including the same proportion of Q&A and statements in both subsamples. We analyzed the sample through LIWC, converted the percentages of each category into raw counts in order to be able to run a bivariate correlation, and screened the 95 variables by selecting those which presented a significant correlation and an effect size of $r > 0.20$. As some of the resulting categories contained many overlapping items and were umbrella categories, we conducted an ulterior screening process by analyzing collinearity and eliminating the broader variables (e.g. function) highly correlating with more specific ones (e.g. personal pronouns, impersonal pronouns). We ran a logistic regression to analyze whether a
model could be delineated that would be able to distinguish the condition (FC or CNID) starting from the three predictors which explained the greater amount of variance without overfitting the model: personal pronouns, impersonal pronouns, and conjunctions. In order to cross-validate our results, we created a second random sample of 12 FC and 12 CNID with the same format proportions, and the predictors identified in the first sample were tested on this out-of-model sample. In order to understand which specific words weighted the most on the predictors, as they correspond to dictionary categories containing multiple words, the frequency of the specific words within the categories were researched in the texts and the eleven most affluent terms were identified for each category. Lastly, to understand in the context in which these words were used and what linguistic function they performed within the corpus of interest, we conducted a qualitative analysis of concordances through AntConc.

Results

Baseline Characteristics of the Corpora

AntConc revealed that the FC corpus consisted of 162,284 words, 3.6% of which were unique (5,895), while in the CNID corpus, 4.9% (5,740) of the 116,024 words were unique. We conducted a bivariate correlation analysis on the single files output generated by LIWC after converting the percentages into raw counts. Reported below are the categories that showed a significant degree of correlation ($p < 0.05$, $r > 0.2$) with either FC (if they had a negative $r$ coefficient, as we coded FC as “0”), or CNID (if they had a positive $r$ coefficient, as we coded CNID as “1”).

FC were correlated with the analytic category ($r = -0.366$, $p < .001$), suggesting that the speech is more formal, logical and hierarchical, as opposed to informal and narrative. They were also correlated to the clout category ($r = -0.219$, $p = .011$), pointing toward a confident speaker who perceives a high level of expertise regarding the topic being discussed, possibly indicating higher social status, confidence and leadership (Kacewicz, Pennebaker, Davis, Jeon,
& Graesser, 2013). Other categories that defined the FC corpus were *they* \((r = -.377, p < .001)\) which is characterized by words related to the third person plural such as “them” or “their”, *ipron* \((r = -.307, p < .001)\), consisting of impersonal pronouns like “it” and demonstrative pronouns like “this” or “those”, *assent* \((r = -.208, p = .016)\), which includes terms of agreement such as “okay” or “yeah”, and *nonflu* \((r = -.266, p = .002)\), which is characterized by discourse fillers and hesitators (e.g. hm, oh, ahh). In summary, false confessors’ speech seems to be characterized by impersonal pronouns, terms of agreement, hesitators, a high level of confidence, and formal, logical language.

CNID were instead correlated with the category named *Sixltr* \((r = .203, p = .019)\) consisting of words longer than six letters and used as a proxy for word complexity. The strongest correlation was found with the category *function* \((r = .365, p < .001)\), which is comprised of many subdimensions, including personal pronouns, articles and auxiliary verbs. As personal pronouns (*ppron*), and auxiliary verbs (*auxverb*) also stood out as single categories (respectively \(r = .207, p = .016; r = .241, p = .005\)), it is likely that the significance of *function* is mostly influenced by these two categories. The category *I* was also found relevant \((r = .278, p = .001)\) and since it is contained within *ppron* and therefore *function*, it suggests that in CNID speakers have a higher use of the first-person pronoun, along with its possessive forms and declinations. CNID also correlated with the categories *conj* \((r = .220, p = .011)\) representing conjunctions such as “and”, “but” or “because”, *adj* \((r = .232, p = .007)\) consisting of adjectives, and *drives* \((r = .210, p = .015)\), which is a general category including words associated with subcategories: affiliation, achievement, power, risk and reward. It is to be noted that none of these subcategories turned out to be significant, although *risk* reached a \(p\)-value of .055, and could partly explain the significance of the broader *drives*. Taking the topic of discussion into consideration, it is not surprising to find *risk* words associated to confessions statements, as its vocabulary includes words conveying loss and danger that would be expected in high stakes
situations. Yet, its correlation with CNID as opposed to FC suggests further research should explore this pattern. Albeit non-significant, the LIWC category labeled as tone, which corresponds to the emotional state of the speaker and suggests that scores below 50 are to be considered signs of negative emotions, varied between conditions. False confessors overall scored 10 percentage points higher (30.88) than confessors not in dispute (20.70), possibly indicating the turmoil caused by the knowledge of having indeed committed the crime of the latter.

**Differences Between the Corpora**

To test whether any variables or combination of variables would be able to discriminate between the two types of confessions, we conducted logistic regression on a random sample of FC \( n=25 \) and a CNID \( n=25 \) containing the same proportion of statements and Q&A to control for format. As LIWC’s output offers 95 variables, we ran a bivariate correlation analysis to identify variables which showed a significant degree of correlation with FC and CNID \( p < 0.05 \) and \( r > 0.2 \), and obtained the following predictors: analytic, clout, function, pronoun, ppron, I, they, ipron, conj, female, comma. As some of LIWC’s categories contain overlapping words, we generated a correlation table of these predictors to assess collinearity (the full correlation matrix can be found in the Appendix A). If two or more categories showed a high degree of collinearity, we chose the variable with the highest level of specificity over the broader one. Through this process, we eliminated the variables analytic, clout, function and pronoun, as they are umbrella variables containing the other ones. We then tested the remaining seven categories in different combinations to find out which ones would explain the highest amount of variance without overfitting the model. We found that comma and female did not add predictive power to the model, and hence we eliminated them. Lastly, as I and they belong to the ipron and ppron categories, and we found the latter to be more predictive than the single two pronouns, we chose to keep the category ipron and ppron over I and they. In summary, the
three variables that we found to explain the greater amount of variance without overfitting the model were: impersonal pronouns (e.g. it, what, that, stuff, etc.), personal pronouns (I, he, she, him, us, their etc.), and conjunctions (and, but, because, cuz, then etc.).

The overall model was predictive of the dependent variables of FC or CNID ($\chi^2=22.439$, $df=3$, $p<.001$), and the three variables were able to correctly classify 37 of the 50 confessions (Table 1). Impersonal pronouns had a significant association with FC ($B=-44.974$, $e^B<.001$), coded as 0, while CNID were coded as 1, therefore explaining the negative coefficient accompanying this variable. CNID, instead, presented an association with personal pronouns ($B=.448$, $e^B=1.566$) and conjunctions ($B=8.508$, $e^B=4958.482$). These results suggest that FC are characterized by a higher degree of impersonal pronouns while CNID contain more conjunctions and personal pronouns. In short, this model was able to discriminate between the confessions that were false vs not in dispute with an overall accuracy rate of 74% (Table 2).

We further tested our predictors on an out-of-model sample composed of 12 FC and 12 CNID containing the same proportion of Q&A and statements as Model 1. Model 2 was also predictive ($\chi^2=22.309$, $df=3$, $p<.001$) and the direction of the variables replicated our previous results, associating impersonal pronouns to FC, and personal pronouns as well as conjunctions to CNID (Table 1). The accuracy of this model was also corroborated by the classification rate of 83.3% across both conditions, as it was able to correctly classify 20 confessions out of 24 (83.3%; see Table 3). In summary, our model suggests that false confessors use more impersonal pronouns, such as it, that, what, thing etc., and less personal pronouns, like I, he, she, them, us, etc., which instead are more likely to be more found in CNID. Conjunctions such as and, but, because, or etc., were also more frequently used in CNID.

**Predictors: Linguistic Function and Context**

Impersonal pronouns, personal pronouns and conjunctions represent three linguistic categories containing a multitude of lemmas that during the development of LIWC were
associated with the broader concept of the category by independent raters. In order to understand which words within each category influenced these results, we created frequency table to assess how many times each word appeared in the corpora (Figures 1-3). As displayed by Figures 2 and 3, the words “I” and “and” were responsible for the majority of the hits within the personal pronouns and conjunctions categories. “And” in particular was far more frequent in the CNID corpus, at a little over 4% while only 2.4% within FC. Chafe (1993) analyzed the frequency of the most used conjunctions and found that the word “and” comprises approximately 4.4% of all spoken English, while being only 1% of written language. Comparing these statistics to the ones emerged from our sample, the frequency of “and” within CNID appears to fall within what is expected for spoken language, while its rate in FC suggests a closer proximity with written language.

Within the impersonal pronoun category, “it” and “that” were the most frequent impersonal pronoun used in the FC corpus (Figure 1); it is interesting to notice that even though impersonal pronouns correlated with FC, the first two most frequent words were still less frequent among FC than in CNID, while it appears that the less frequent words were the ones that influenced the final association between this category and FC.

Subsequently, we conducted an in-depth qualitative analysis through AntConc to understand the functions and context of “I,” as it was the most frequent pronoun in both corpora. The concordance analysis revealed that the pronoun “I” in CNID was paired more often with lexical verbs, which express action and state, and include all verbs but auxiliary ones (e.g. went, took, pulled, found) (Biber et al., 1999). A subset of lexical verbs concerning physical violence was further analyzed and words like hit, cut, killed, raped, shot etc. were found to be seven times more prominent in CNID than in FC. We also identified a disparity within the category of mental verbs, which express cognition (e.g. think, notice, realize, know), as the most frequent mental verb within CNID was “I think/thought”, while within FC the first
rank was occupied by “I know/knew”. Moreover, FC displayed higher proportions of “I guess” and “I mean”, while almost all the other cognitive verbs frequently identified in the corpora, including all dynamic (e.g. decided, realized, noticed) and most sensory (e.g. heard, looked, saw/see) verbs appeared more frequently in the CNID, with the only exception of feel/felt, which appeared more in FC. Lastly, as a difference was observed in the frequency with which FC and CNID used the sentence “I do not/don’t remember/recall,” the context surrounding it was examined. The sentence appeared 57 times in FC, spread over 13 confessions, while in the CNID corpus it was uttered 86 times over 36 confessions. When the words preceding and following it were inspected, a pattern emerged: in FC, 49% of the total utterances ended the sentence and the speaker did not expand on what was not remembered, while in CNID, only 19% of the time the words terminated the sentence, while in the remaining 81% the speaker explained what was not remembered (e.g. “I don’t remember what he was saying”, “I don’t remember if I had blood on them”).

Discussion

This study analyzed the linguistic differences between FC and CNID, based not on content but on language style. First, I delineated the linguistic characteristics of both types of confessions, starting from percentage of unique words over total words, with CNID displaying a higher percentage (4.9%) over FC (3.6%), thus pointing towards a higher lexical variety in CNID than in FC. The language of FC in our sample was characterized by impersonal pronouns, formality, and logical language suggesting a confident speaker, as well as a higher number of hesitators and terms of agreement. These last two categories were not considered significant, as the transcriptions of the confessions did not always include hesitators (e.g. ahh, mmh, or ellipses) and the high frequency of terms of agreement depended on the preponderance of the Q&A format in FC. In contrast, CNID were characterized by a higher frequency of words longer than six letters, suggesting greater language complexity, more personal pronouns,
especially “I”, and auxiliary verbs, as well as conjunctions, adjectives and possibly words regarding risk. The prevalence of words longer than six letters in CNID, and therefore the corresponding lack thereof in FC, could be linked to demographics and cognitive variables. In fact, we know that juveniles and individuals affected by a mental illness are more susceptible to producing a false confession (Drizin & Leo, 2004; Redlich, 2004; Redlich, Silvermann, Chen, & Steiner, 2004) and this could explain the lower lexical variety.

Three linguistic predictors we identified were able to discriminate between FC and CNID with an accuracy rate between 74% and 83.3%: impersonal pronouns, personal pronouns and conjunctions. A higher frequency of impersonal pronouns was associated to a higher likelihood of a confession being false, while confessions containing more personal pronouns and conjunctions were associated with a higher likelihood that the confession belong to the sample not in dispute. The lesser usage of first-person singular by false confessors could be tied to a study which associated this pattern with deception, as a way to distance oneself (Newman, Pennebaker, Berry, & Richards, 2003). Moreover, as the authorship of FC may not be as clear as the one of the CNID, and the resulting confession may be due to a dynamic contamination of the speech between suspect and law enforcement, it is of interest that a study by Hancock, Curry, Goorha, and Woodworth (2008) found that both participants who lied and were lied to showed the same linguistic markers: lower word count, less first-person singular and more sensory words.

A qualitative analysis of “I” also revealed a pattern of association with lexical verbs (e.g. put, took, killed, hit etc.) in CNID, and the lack thereof in FC. Following are some examples of lexical verbs within the CNID corpus:

- “[...] and then eventually I put everything in the closet.”
- “I hit her in the side of the neck with my right forearm and she fell off the bed.”
- “I raped her, it was all a drug induced cloud.”
Moreover, the most frequent mental verb following “I” in FC was found to be “I know/knew”, while for CNID it was “I think/I thought”, further raising the question of authorship, as it also ties to the higher level of confidence found in FC by LIWC. However, both “I guess” and “I mean” were more frequently used in FC, adding a tentative aspect to the narrative and contrasting the previous finding, perhaps suggesting partial authorship. List below are examples of the usage of “I guess” and “I mean” in FC:

- “I guess I was, yeah, I was kind of drunk by then, drinking pretty much.”
- “So, I guess Emilio took her in the bedroom, in his bedroom, and started, I don’t know, he beat her up, I guess. He knocked her out of something like that; he said in order to make love to her; she started yelling again and started fighting him, I guess he beat her up real bad and Ray got pissed about it.”
- “And I mean so I... that’s why I got that gun for that purpose. But I don’t... I mean I don’t need no gun you know what I mean.”
- “But the problem is, that’s life, life goes on. Life goes on. Life goes on. I mean myself, I mean...”

Lastly, a close examination of the variations of the sentence “I don’t remember” revealed that in FC it tended to complete the sentence in approximately half of the cases. The following are instances of the use of the words “I don’t remember” in FC:

- “I don’t remember, I don’t think, I don’t know nothing.”
- “Not that I remember, I don’t remember nothing, I don’t remember nothing.”
- “I was – don’t – don’t remember. I was drinking that night.”
- “I don’t remember exactly.”

In contrast, in CNID, the sentence was followed in the majority of the cases by an explanation of what was not remembered. These are some decontextualized examples of how this sentence was used in CNID:
• “I don’t remember how many times I stabbed him.”

• “I guess, I don’t remember if it was on the bed or the floor.”

• “I don’t remember the exact words I used, but I told him what happened and where it was.”

• “I don’t remember if she was in or out of the car when she asked.”

This discrepancy could be due to the actual lack of knowledge of FC in regards to the facts of the crime, therefore their “I don’t remember” would symbolize a more general absence of memory caused by absence of the fact, while for CNID the context shifts toward a forgetfulness of specific details of the crime. This difference could also be tied to the cognitive processes involved in the production of images rather than false memories, when it comes to FC. False memories involve the actual belief of having experienced the remembered event, while images are conceptualized as associated with the suggested event but not experienced as memories of the event (Lindsay et al., 2004; Desjardins & Scoboria, 2007; Hessen, Kayfitz & Scoboria, 2012). Another study further distinguished the two by stating that people generating false memories “claimed to remember the event and reported at least two specific details about it,” while individuals who experience images only “speculated about at least three different aspects of the event” (Strange, Wade, & Hayne, 2008, p. 479). Thus, false confessors may be generating images rather than false memories in the majority of the cases, and that could be the reason behind the lack of detail following the unremembered events. This would also explain the discrepancies we found among mental verbs between FC and CNID, with the latter containing more dynamic and sensory verbs, which are used in narration to express agency (decided, realized, noticed) and the sensory experiences of the narrator, therefore actively placing the individual in the memory.

Knowing how pronouns are used in different types of confessions is an important factor in the determination of the nature of a confession, and this model could be used in the future to
assess whether a confession could be true or false, within some probabilistic level of certainty. However, it is yet unknown why the speech of FC turned out to be different from that of CNID. A possible explanation could consist in different speech patterns between innocent confessors and guilty confessors, or it could be due to the different type of pressure these two populations find themselves under.

Another plausible explanation instead involves differences in the authorship of the confession. History presents all too many instances in which police appeared to write an innocent person’s confession. In 1963, New York City detectives questioned George Whitmore, a 19-year-old African American man, for 26 hours, which produced a detailed 61-page confession to two high-profile murders. Whitmore was ultimately exonerated. His false confession, however, plainly authored by police, was so troubling that in *Miranda v. Arizona* (1966), the U.S. Supreme Court cited Whitmore as a “conspicuous” example of police coercion in the interrogation room (see Kassin, 2017; this infamous case is fully described by English, 2011; Shapiro, 1969). With 95% of false confessions containing accurate crime facts known to police that the innocent suspect could not have known, it is possible that in the case of FC there may have been a heavier contamination of police speech into the suspect’s speech, or even more directly some of the confessions may have been authored by the officers. Yet, it is to be noted that also among the sample of CNID some confessions were transcribed by police officers upon request of the suspect, or because of departmental regulations on confession evidence, therefore a base level of contamination was to be expected in both samples. However, the different degree in which such contamination happened in FC and CNID may be responsible for the observed stylistic differences between them.

**Limitations and Future Research**

Computerized linguistic analysis software and psychological language analysis programs are still in a developmental stage and they do not factor in context or irony, therefore
some of the context within which the words we analyzed were used may have been misinterpreted due to this limitation. For example, the sentence “I am crying of laughter” would add points to both the positive and negative emotions categories. However, we performed a context analysis on the main predictors to partly obviate this shortcoming. One of the limitations of this study lies in the archival nature of the data, which implies an impossibility of knowing under which circumstances each confession was taken, as well as the level of accuracy of the transcriptions, as video or audio recordings were not present in the files. Interrogations and the process leading to a confession, which is often the result of the first, are currently a black box, unknown to both the public and the judicial system. The final confession is the only evidence revealed during legal proceedings, while often there is not a record of what preceded it. Because of the incredible weight that confession evidence holds, it is of paramount importance for the interrogation process to be video and audio recorded—from start to finish. Another limitation is due to the exploratory and corroborating nature of this study, as we initially tested the model on the data we derived it from, although we also tested it on an out-of-model sample achieving similar results.

Future research should follow up on the context and functions of “it”, as we found this word to be the most frequent within the impersonal pronouns’ category, and also which words and sentences are conjoined by “and”, which was the most common conjunction in our corpora, through the analysis of concordances and collocates to achieve a better understanding of the context in which these words are used. Future research should also investigate style contamination during the interrogation process, and the injection of law enforcement speech in the suspect’s narrative, as well as possible police authorship. More research needs to be conducted also to determine whether there is a stylistic difference between innocent and guilty individuals’ speeches. For example, the cheating paradigm could be used in a controlled experiment to induce participants to give a true or false narrative confession after an
interrogation, therefore allowing us to control for all factors and independently manipulate the variables: guilty/innocence, and true/false confessions. The recordings would then be accurately transcribed in a standardized way, making the linguistic analyses more robust.

In conclusion, this study delineated the stylistic baselines of false confessions and confessions not in dispute and showed how corpus linguistics can help distinguish between the two. As we have now identified three predictors and have two samples with set odds ratios, in the future it would be possible to discriminate other confessions, with a degree of accuracy between 74% and 83.3%, by running a confession through LIWC and then incorporating its output in the preexisting model by substituting it to either a known false confession or a known confession not in dispute from one of the samples, and then analyzing the changes in the classification table.
References

18 U.S. Code § 3501


Dassey v. Dittmann., 570 F.3d 1096 (9th Cir. 2009).


Table 1

Logistic regression of impersonal pronouns, personal pronouns and conjunctions on FC and CNID

<table>
<thead>
<tr>
<th>Variables</th>
<th>Model 1 (n=50)</th>
<th>Model 2 (n=24)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Odds Ratio</td>
</tr>
<tr>
<td>Impersonal pronouns</td>
<td>-.44974</td>
<td>.000</td>
</tr>
<tr>
<td>Personal pronouns</td>
<td>.448</td>
<td>1.566</td>
</tr>
<tr>
<td>Conjunctions</td>
<td>8.508</td>
<td>4958.482</td>
</tr>
</tbody>
</table>

Nagelkerke pseudo r-square  .482  .807

Chi-square       22.439, df=3, p < .001  22.309, df=3, p < .001

Notes: FC had been assigned a value of 0, while CNID had been assigned a value of 1.
Table 2

*Classification table for Model 1 (n=50)*

<table>
<thead>
<tr>
<th>Observed</th>
<th>Predicted</th>
<th>% correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>FC</td>
<td>19</td>
<td>6</td>
</tr>
<tr>
<td>CNID</td>
<td>7</td>
<td>18</td>
</tr>
</tbody>
</table>

Overall percentage 74%
Table 3

*Classification table for Model 2 (n=24)*

<table>
<thead>
<tr>
<th>Observed</th>
<th>Predicted</th>
<th>% correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>FC</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>CNID</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Overall percentage</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 1

*Frequency distribution of impersonal pronouns*
Figure 2

*Frequency distribution of conjunctions*
Figure 3

*Frequency distribution of personal pronouns*
Correlation matrix of significant LIWC categories

<table>
<thead>
<tr>
<th>LIWC Category</th>
<th>Analytic</th>
<th>Clue</th>
<th>function</th>
<th>possess</th>
<th>pposs</th>
<th>i</th>
<th>they</th>
<th>iposs</th>
<th>conj</th>
<th>female</th>
<th>Common</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PC vs CNID</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence Correlation</td>
<td>1</td>
<td>-316*</td>
<td>-299*</td>
<td>.517*</td>
<td>.445*</td>
<td>.333*</td>
<td>.377*</td>
<td>.371*</td>
<td>.333*</td>
<td>.338*</td>
<td>.345*</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.035</td>
<td>.070</td>
<td>.000</td>
<td>.001</td>
<td>.018</td>
<td>.007</td>
<td>.035</td>
<td>.019</td>
<td>.016</td>
<td>.023</td>
<td>.044</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>Analytic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence Correlation</td>
<td>-316*</td>
<td>1</td>
<td>.020</td>
<td>-613*</td>
<td>-436*</td>
<td>-327*</td>
<td>-306*</td>
<td>-304*</td>
<td>-132</td>
<td>-196</td>
<td>-056</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.035</td>
<td>.089</td>
<td>.000</td>
<td>.002</td>
<td>.023</td>
<td>.031</td>
<td>.046</td>
<td>.031</td>
<td>.167</td>
<td>.300</td>
<td>.273</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>Clue</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence Correlation</td>
<td>-299*</td>
<td>.030</td>
<td>1</td>
<td>-305*</td>
<td>-593*</td>
<td>-498*</td>
<td>-580*</td>
<td>153</td>
<td>-076</td>
<td>-338*</td>
<td>-314*</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.050</td>
<td>.089</td>
<td>.020</td>
<td>.000</td>
<td>.003</td>
<td>.000</td>
<td>.090</td>
<td>.009</td>
<td>.000</td>
<td>.007</td>
<td>.096</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence Correlation</td>
<td>.517*</td>
<td>-613*</td>
<td>-316*</td>
<td>1</td>
<td>.465*</td>
<td>.370*</td>
<td>.347*</td>
<td>.305*</td>
<td>.375*</td>
<td>.550*</td>
<td>.259</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.000</td>
<td>.038</td>
<td>.000</td>
<td>.008</td>
<td>.015</td>
<td>.010</td>
<td>.000</td>
<td>.000</td>
<td>.069</td>
<td>.277</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>possess</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence Correlation</td>
<td>.445*</td>
<td>-436*</td>
<td>-593*</td>
<td>.691*</td>
<td>1</td>
<td>.387*</td>
<td>.744*</td>
<td>.062</td>
<td>.357</td>
<td>.418*</td>
<td>.454*</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.001</td>
<td>.002</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.003</td>
<td>.001</td>
<td>.001</td>
<td>.006</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>pposs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence Correlation</td>
<td>.333*</td>
<td>-327*</td>
<td>-436*</td>
<td>.370*</td>
<td>.367*</td>
<td>1</td>
<td>.189*</td>
<td>.114</td>
<td>.152</td>
<td>.140</td>
<td>.327*</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.019</td>
<td>.011</td>
<td>.001</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.043</td>
<td>.031</td>
<td>.033</td>
<td>.020</td>
<td>.016</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>i</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence Correlation</td>
<td>.377*</td>
<td>-300*</td>
<td>-330*</td>
<td>.347*</td>
<td>.344*</td>
<td>.389*</td>
<td>1</td>
<td>.277</td>
<td>.271</td>
<td>.254</td>
<td>.382</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.007</td>
<td>.031</td>
<td>.000</td>
<td>.015</td>
<td>.000</td>
<td>.000</td>
<td>.051</td>
<td>.057</td>
<td>.070</td>
<td>.047</td>
<td>.035</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>they</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence Correlation</td>
<td>-316*</td>
<td>-047*</td>
<td>.133*</td>
<td>-035*</td>
<td>-052*</td>
<td>-104*</td>
<td>-277*</td>
<td>1</td>
<td>.256*</td>
<td>.208</td>
<td>.173</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.035</td>
<td>.076</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.148</td>
<td>.230</td>
<td>.992</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>iposs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence Correlation</td>
<td>-331*</td>
<td>-132*</td>
<td>.076*</td>
<td>-075*</td>
<td>-637*</td>
<td>-152*</td>
<td>-271*</td>
<td>.376*</td>
<td>1</td>
<td>.184</td>
<td>.020</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.019</td>
<td>.061</td>
<td>.000</td>
<td>.065</td>
<td>.094</td>
<td>.017</td>
<td>.000</td>
<td>.000</td>
<td>.020</td>
<td>.092</td>
<td>.361</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>conj</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence Correlation</td>
<td>.339*</td>
<td>-198*</td>
<td>-331*</td>
<td>.599*</td>
<td>.488*</td>
<td>.140</td>
<td>.034</td>
<td>.208</td>
<td>.184</td>
<td>1</td>
<td>.310</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.016</td>
<td>.087</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.140</td>
<td>.300</td>
<td>.876</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>female</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.002</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>Common</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presence Correlation</td>
<td>-356*</td>
<td>-139*</td>
<td>-086*</td>
<td>-174*</td>
<td>-160*</td>
<td>-101*</td>
<td>-090*</td>
<td>.001</td>
<td>.132</td>
<td>.253</td>
<td>.022</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.044</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>