*Article*

# Speech Characteristics as Indicators of Personality Traits

**Sinae Lee [1],\*, Jangwoon Park [2] and Dugan Um [2]**

[1]   Department of English, Texas A&M University—Corpus Christi, Corpus Christi, TX 78412, USA

[2]   Department of Engineering, Texas A&M University—Corpus Christi, Corpus Christi, TX 78412, USA;
    jangwoon.park@tamucc.edu (J.P.); dugan.um@tamucc.edu (D.U.)

\*   Correspondence: sinae.lee@tamucc.edu

**Abstract:** This study examines the relationship between speech characteristics and personality traits by drawing on pseudo-naturalistic conversations and on personality dimensions identified by the Myers–Briggs Type Indicator (MBTI) model which assesses four personality dimensions of introversion-extroversion, sensing-intuiting, thinking-feeling, and judging-perceiving. The speech of 30 participants was recorded and transcribed, after which a number of speech features including pitch, loudness, response time (i.e., how fast one responds to a prompt), speech rate, and discourse markers were extracted and analyzed. Results show that several speech features correspond to different personality dimensions. Specifically, speech rate as measured by words per minute reveals significant differences between judging individuals and perceiving individuals (perceiving individuals speak faster than judging individuals); there is a significant difference in response time for extroverts and introverts (extroverts respond faster); a significant difference is observed in loudness between judging and perceiving individuals (judging individuals are louder). The frequency of discourse markers is significantly higher for intuiting individuals than sensing individuals. The study discusses these findings in further inquiring the relationship between language and personality.

**Keywords:** personality traits; Myers–Briggs Type Indicator (MBTI); speech characteristics; naturalistic conversation; vocal features

## 1. Introduction

Language can be a useful window through which we can understand the social and psychological attributes of its users, including personality traits. The utility of language in identifying personality has long been recognized and investigated [1–7], with many recent studies observing correlations between specific language features and certain personality attributes. For example, Jurafsky et al. (2009) [8] and Ranganath et al. (2009) [9] showed that linguistic and interactional features such as speech rate, pitch range, and laughter are closely related to friendliness or flirtatiousness, to an extent that these features can be used in detecting the speaker's intent to be friendly or to be flirtatious at above 70% accuracy. Weninger et al. (2012) [10] identified a number of acoustic features such as loudness or voice quality that can be used to recognize speakers that exhibit personality characteristics of leadership (e.g., high-achieving, charismatic, team-playing, etc.). Holtzman et al. (2019) [11] observed that certain linguistic items, such as sports-related words, swear words, and second-person pronouns, may mirror narcissistic personality.

Studies have also examined how language reflects personality dimensions identified by the Myers–Briggs Type Indicator (MBTI) model. The MBTI model measures four personality dimensions of introversion–extroversion, sensing–intuiting, thinking–feeling, and judging–perceiving, and attempts to describe an individual's personality as one of 16 types. The MBTI model utilizes a series of questions that are designed to gauge a person's inclinations for outwardness, sociability, reliance on senses, orderliness, and so forth. Respondents to the MBTI questionnaire answer a series of questions on a Likert scale (e.g., very unlikely to very likely). Based on the compiled answers, the model proposes

the likelihood of the person's introvertedness (compared to extrovertedness), the degree of sensing (compared to intuiting), that of thinking (compared to feeling), and that of judging (compared to perceiving). The introverted/extroverted dimension reflects a person's orientation toward the outer world. The sensing/intuiting dimension differentiates individuals based on how memory and information is processed and stored; a prototypically sensing individual processes information by relying on physical, sensory vehicles, but a prototypically intuiting individual places a subjectively determined interpretative lens when processing information. The thinking/feeling dimension reflects one's style of interaction or decision-making; "thinking" individuals are more logical, task-oriented, and less compromising in prioritizing values, whereas "feeling" individuals are more flexible and tactful in making decisions, actively taking into account different contexts and situations. The judging/perceiving dimension groups individuals based on whether they appear more structured and orderly in carrying out daily tasks; the judging types are thought to be more orderly, while the perceiving types are thought to be more open in selecting and prioritizing what is to be done next. Of these different personality dimensions, the most commonly investigated dimension is extroversion, though there are studies that explore the relationship between language and other personality dimensions in addition to extroversion, such as conscientiousness, agreeableness, neuroticism, or openness [7,12–14]. Studies have identified a number of linguistic characteristics that are associated with extroversion; it has been found that extroverted (or "outgoing") Dutch speakers tend to use linguistic markers of abstraction more frequently than introverted Dutch speakers [15], and also that extroverts tend to be more verbose than introverts [16–20]. Relatedly, extroverts were observed to exhibit higher verbal fluency not only in their own speech [21], but also during interaction as manifested with their shorter response time [22,23].

These linguistic features identified as markers of a particular personality dimension can also be used in predicting personality. This has been shown by Mairesse et al. (2007) [24], who included a number of language features (e.g., thematic word groups, utterance types, pitch, verbosity measured by voice time compared to silent pause, speech rate measured by word per second, etc.) that are related to certain personality factors in building a predictive model, and achieved up to 73% classification accuracy for predicting certain personality traits such as extroversion or openness for experience. Guidi et al. (2019) [25] found the significant correlations between personality traits and speech features, including pitch and voice quality. Particularly focusing on acoustic and vocal features, Mohammadi and Vinciarelli (2012) [26] and Polzehl et al. (2010) [27] similarly showed that the personality traits can be predicted by pitch range, intensity, loudness, formants (acoustic energy that reflects resonance depending on the shape and size of the human vocal tract), spectral analysis (the distribution of acoustic energy across frequency), or speech rate.

The current study is an attempt to continue this line of inquiry by exploring the relationship between speech and personality. Drawing on speech produced during pseudo-naturalistic conversations, we extracted linguistic information from four domains of language use; namely, the time taken in responding to a conversational prompt (i.e., response time) extracted from the interactional domain, speech rate from the paralinguistic domain, pitch and intensity from the acoustic domain, and discourse markers from the text domain. Based on previous studies, we hypothesize that these speech characteristics are associated with some aspects of personality traits as identified in the MBTI model, such as extroversion–introversion, sensing–intuiting, thinking–feeling, or judging–perceiving. Thus, the extracted speech characteristics were examined for their associations with the aforementioned personality traits. In addition, we included gender as part of our variable in examining the associations between speech characteristics and personality traits. This is to ensure that we observe the potential relations between speech and personality on variables that exhibit different ranges depending on the gender of participants, such as pitch. Furthermore, we look at whether men and women, in general, display different patterns for the examined speech characteristics, as part of our effort to search for potential future directions.

## 2. Methods

### 2.1. Participants

We collected semi-natural conversations of 30 individuals (12 males and 18 females) from South Texas. These conversations were collected between 2019 and 2021, and all participants were native speakers of English. The average age of the participants is 32.8 years, and age ranges from 19 to 66. The conversations took place in the frame of sociolinguistic interview [28], a speech elicitation technique in which the interviewer's sole goal is to elicit spontaneous and natural talk in large quantities. To achieve this, a typical sociolinguistic interview is often lengthy in duration, spanning from 30 min to 90 min or so, and loosely structured; most interviewees become more relaxed and less self-conscious as the interview continues, which in turn yields more naturalistic speech. Prior to collecting interviews, the study protocol was approved by the Institutional Review Board (IRB) at Texas A&M University—Corpus Christi.

### 2.2. Experimental Protocol

Two interviewers were recruited and trained to conduct sociolinguistic interviews. The interviewers were undergraduate students, and were native speakers of English. We recruited volunteers by contacting individuals that would potentially be interested in participating in the study. Some participants were students, but not all. All interviews were conducted one-on-one; in other words, one interviewer and one participant were involved in each interview. Each interview was approximately 40 min long on average, and the topics covered during the interviews include family/friends, neighborhood change, school years, memorable incidents, etc. Though the interviews were loosely structured, they were standardized across participants. Given the nature of sociolinguistic interviews, in which the main goal of an interviewer is to elicit the most naturalistic speech, the topics are not strictly controlled; whenever possible, the interviewee is given a full control of the speech event, and when a particular topic is conducive to eliciting much talk, the interviewer dwells on this topic as long as possible. The interviews were audio-recorded, using portable solid-state recorders (Zoom H2n Handy Recorder, Zoom Corporation, Tokyo, Japan). Following the interview, participants were requested to fill out the MBTI test survey.

### 2.3. Data Segmentation

All recoded voice data, without cutting or manipulating any specific parts, was used to analysis. Note that, because we used naturalistic conversation data, the length of the conversation measured from each participant vary depending on the topic of the conversation or questions from the interviewer. Each interview recording was manually segmented based on the percept of a "breath group", which often coincides with an intonational phrase beginning with a louder volume and/or higher pitch ending with a smaller volume and/or lower pitch. The segmented chunks were then manually transcribed as heard, reflecting verbatim production of the utterance, which includes false starts, stammers, fillers, errors, curses, etc. Along with the speech of the interviewees, the speech of the interviewers was transcribed as well. The utterance-level segmentation and transcription were performed using a transcribing software ELAN (Version 6.0) [29]. The audio recordings and corresponding transcripts were inputted into FAVE (Forced Alignment and Vowel Extraction) [30], a computational tool that processes a recording of speech and its word level transcript and generates a file that contains segmentation of not only words, but also units within a word, such as vowels and consonants.

The generated file is in the format of TextGrid, which is essentially a text file object that Praat—a speech analysis software used for this study—can generate in order to label or annotate the audio file [31]. Once generated, a TextGrid file enables us to work with the segments that are aligned with the corresponding acoustic signals. An illustration of a Praat window where both the acoustic signals (soundwave and spectrogram) and the TextGrid are loaded is provided in Figure 1.
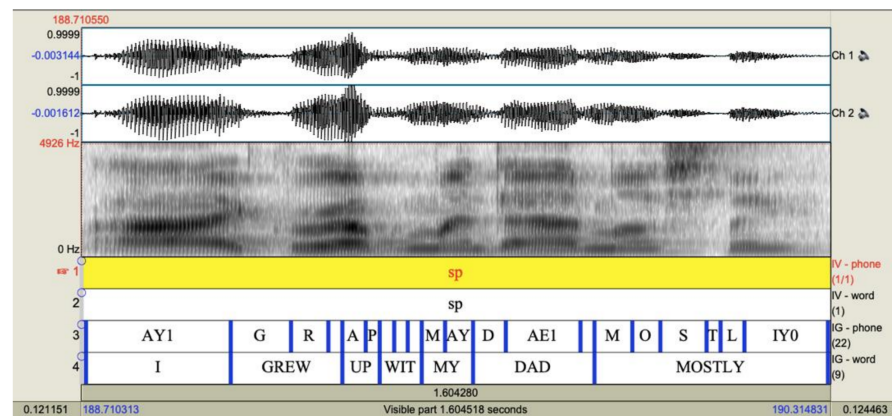
**Figure 1.** An illustration of the Praat window ("sp" refers to "silent pause"; Ch1 and Ch2 refer to Channel 1 and Channel 2, respectively, which indicate the stereo mode of recording; IV, short for Interviewer; IG, short for "Isiah Goodson", which is the pseudonym of this participant).

To sum up, the final dataset included the audio files and the corresponding TextGrid files (containing multi-level annotations including utterance, word, and phone such as vowels and consonants). The data collection and processing procedure is schematically illustrated in Figure 2.
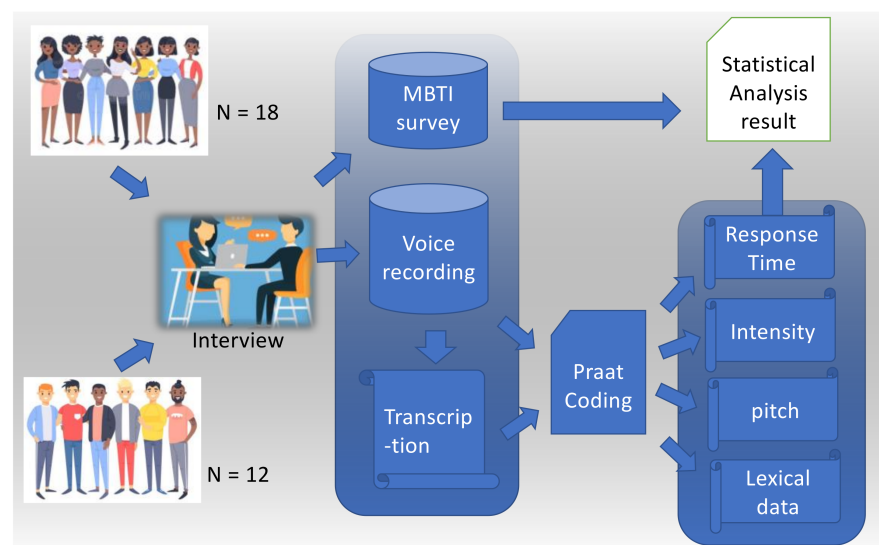


**Figure 2.** Data collection and processing procedure.

### 2.4. Extracted Features

We selected five features that may potentially be associated with some aspect of the personality domains, namely, response time (in milliseconds, or ms), intensity (in dB), pitch (in Hz), speech rate (words per minute, or wpm), and discourse markers. These selections were made either because previous findings on the relationship between language and personality included these features or because we suspected that these features might be utilized when people perceive or judge others' personality. Because these features are drawn from various domains of language use, reflecting interactional attributes (i.e., response time), acoustic attributes (i.e., pitch and intensity), content or textual attributes (i.e., discourse markers), as well as paralinguistic (linguistic-external) speech attributes (i.e., speech rate), examining them would allow us to consider language in its interaction with personality from a more holistic perspective.

For the current study, six features are identified. Table 1 summarizes the descriptions of these features. They are demographic information (gender), and vocal and speech features (response time, voice frequency, loudness, speech rate, and frequency of discourse markers). First of all, gender is considered as a variable that can potentially explain the differences in phonetic and linguistic characteristics of people who have different personality traits. Second, it is presumed that response times during natural conversations could be related to their personality traits based on previous reports [23–32]. Third, voice frequency was selected to investigate whether there is a significant difference in frequency of people's voices among the various personality traits. Fourth, it is assumed that the speech loudness could be significantly different between the personality traits. Fifth, we posited that the talking speed might be associated with personality traits. Lastly, it is hypothesized that the personality traits could be related to the types of words people use–specifically discourse markers–in natural conversations.

**Table 1.** Descriptions of the extracted features to identify the relationships with the personality traits.

| Independent Variables | | Description |
|---|---|---|
| Gender | Male or Female | The Gender of a Participant |
| Response time | Mean response time (ms) | Average time a participant takes in responding to the question or a prompt, in milliseconds. Response time refers to the duration between the end of the interviewer's question or comment and the beginning of the participant's response. |
| Voice frequency | Pitch (Hz) | Average pitch of a participant in Hertz (Hz). |
| Loudness | Mean sound intensity level (dB) | An average magnitude of a participant's vocal energy in decibel (dB) throughout the entire speech. |
| Speech rate | Mean word per minute | An average number of words spoken within a minute by a participant. |
| Frequency of Discourse markers | Types: like, well, uh/um, just, kind of/sort of, I mean, you know | The frequency rate of each discourse marker per participant calculated from dividing the number of occurrences for each marker by the number of total words uttered throughout the interview. |

Response time (RT) refers to the amount of time a speaker takes before responding to a prompt, whether the prompt is provided in the form of a question or a statement. A Praat script was written to process the interviews in our dataset, in which all the instances of RT (i.e., the averaged time between the end of an interviewer's speech and the start of a response) were measured. Each TextGrid file contains four "tiers", whereby the top two tiers (Tiers 1 and 2) are annotations for the interviewer and the bottom two tiers (Tiers 3 and 4) for the interviewee. Tiers 1 and 3 contain the phone-level annotation, and Tiers 2 and 4 contain the word-level annotation (as illustrated in Figure 1). For each interview, the script started by scanning Tier 4, determining whether the interviewee was talking or silent. If silent, where the segment was labeled as "sp", the script then read the annotation in Tier 2 in order to see if it is also marked as "sp". If both parties (the interviewer and the interviewee) were silent, then it was assumed that this silent time constitutes a RT for that particular interactive moment. The script then identified the start time of "sp" by the interviewer as well as the start time of "not sp" by the interviewee following the interviewer's "sp", thereby calculating the duration of the RT in milliseconds. When the script was run toward the end of a recording, all the instances of RT were collected in order to compute average, minimum, and maximum RT for each participant throughout the interview.

In order to measure pitch and intensity, another Praat script was written and run, focusing on the speech produced by the interviewees only. For both the pitch and intensity, the measurement was taken on vowels only (only the stressed vowels were measured). In order to measure pitch, the script extracted 100 pitch values per second for a vowel token, using the automated pitch measurement, specifically the unbiased autocorrelation method developed by Boersma (1993) [32]. For both male and female speakers, the pitch floor was set as 75 Hz and the ceiling set as 600 Hz. This range is set as standard by the automated

pitch measurement provided in Praat. While we acknowledge that the pitch floor is often recommended to be set differently depending on the gender of the speaker [33]—often around 70 Hz for men and 100 Hz for women, we did not differentiate the pitch setting for the speaker gender. There are two reasons for this: Firstly, the adjustment of pitch floor is an especially pressing matter when the background noise is present; however, all of the interviews conducted for this study have minimal background noise. Secondly, many speakers, especially younger female speakers, employ a creaky voice, often co-occurring with a very low pitch. It is often advised to set the pitch floor as low as 50 Hz or even 40 Hz if the recording contains predominantly creaky voice. Given that the creaky voice is fairly common across the participants in our data, we felt that setting the pitch floor of 75 Hz for both males and females was apt in measuring more pitch points per speaker, which in turn would grant us a higher confidence in our calculations for the average pitch per speaker.

Per token, the mean fundamental frequency was computed. Vowel tokens with duration shorter than 50 milliseconds were excluded. In obtaining the degree of loudness, or intensity, the intensity of a sound in air was measured using units of Pascal, and expressed in Decibel (dB). All the measurements for intensity were taken at the midpoint of a vowel. Measurements for the pitch and intensity were compiled and computed for average, minimum, maximum, as well as variance. In examining the pitch difference between the two opposites within a given personality dimension (e.g., introverted vs. extroverted), we separated male participants from female participants.

Speech rate and discourse markers were analyzed using R package "tidytext" [34]. Firstly, each utterance (i.e., intonational phrase) was parsed down to individual words, and the duration of the utterance was also logged. Per utterance, the number of words per minute was calculated as a measure of speech rate; this was conducted by dividing the number of words multiplied by 60 with the duration. All the utterances per participant were compiled and computed for average, minimum, maximum, and variance.

In order to examine the relationship between discourse markers and personality, a total of nine discourse markers—like, well, uh, um, just, kind of, sort of, I mean, and you know—were identified. "Uh" and "um" were treated as one, as well as "kind of" and "sort of". This yields a total of seven types of markers. Per participant, the frequency of each of these markers ((the number of occurrences for each/the total number of words in the interview) × 100) was generated. For two-word markers (i.e., kind of, sort of, I mean, and you know), the frequency rate was calculated in reference to the total number of bigrams—instead of the total number of words—in the interview.

## 2.5. Statistical Analysis

In this study, based on the central limit theorem, we presumed that the extracted variables from the sound recording for 40–60 min per person for a total of 30 subjects follow a normal distribution, especially for the non-lexical features. To determine whether there is a statistical difference of the extracted features between personality types, two sample *t*-tests were performed for each of the four MBTI personality types. Particularly for the personality types in which the number of participants is small, the statistical analysis results would be best if used as a reference. It is also the case that the statistical significance would often be low when considering the small number of sample size. For discourse markers, a series of two-sample *t*-tests was conducted as well in order to see if there were any differences among participants, depending on their gender (male vs. female), and also on their personality types (e.g., introverted vs. extroverted), in terms of the frequency of these markers. Normality tests (the Shapiro–Wilk test) were performed on the frequency of the markers, after which the type of statistical test was determined. If the hypothesis of normality is rejected, the Wilcoxon rank-sum test was performed. For almost all markers (with an exception for "all markers combined" and for "like"), the distribution was right-tailed (i.e., positive skew). Nevertheless, an important part in this study is to compare the magnitude of differences in the extracted features between personality types. In the results section, two tables show the average difference in sound and linguistic characteristics by

gender and MBTI personality traits. More in-depth interpretations are discussed in the discussion section.

## 3. Results

### 3.1. Characteristics of Non-Lexical Speech

Table 2 shows statistical testing results on each non-lexical characteristic of the speech data, namely, pitch, intensity, response time, and speech rate. As a result of performing the two-sample t-test, it was found that the difference between male and female was significant on pitch ($p < 0.001$), reflecting the physiological differences between the two sexes. For the females, the average pitch was 189.6 Hz (SD = 21.0, minimum = 153.7 Hz, and maximum = 229.3 Hz), and for the males, the average pitch was 134.6 Hz (SD = 12.4, minimum = 121.6 Hz, and maximum = 163.2 Hz). The average pitch of the females that were classified as "intuiting" was approximately 18.8 Hz higher than the sensing females, though this does not safely reach a statistical significance ($p = 0.056$).

During the interview, the interviewers asked a number of questions that likely prompted answers of varying length; some answers were as brief as one word (e.g., "yes", "no", etc.), while others were as long as 190 consecutive utterances (an utterance was operationalized as an intonational phrase). On average, each interviewee produced 9.45 utterances in response to a question (SD = 18.1). There was no significant difference between men and women in the speech rate measured by words per minute. However, there was a significant difference in the speech rate between judging and perceiving groups ($p = 0.046$). The participants that were classified as "judging" uttered fewer words in a minute, on average, than the "perceiving" participants (186 and 202 words per minute for the "judging" and "perceiving" participants, respectively). In terms of response time (unit: ms), there was no gender difference, nor was there any meaningful difference between sensing individuals and intuiting individuals. However, the extroverts responded to the prompts about 0.329 s faster, on average, than the introverts ($p = 0.035$). Lastly, in terms of intensity (unit: dB), there was no significant difference between the males and females. The introverted/extroverted dimension, sensing/intuiting dimension, and thinking/feeling dimension showed no effect on the degree of intensity, either. However, the participants that were classified as "judging" were found to be about 4.1 dB louder, on average, than the "perceiving" participants ($p = 0.009$).

### 3.2. Characteristics of Discourse Markers

Table 3 summarizes statistical testing results on the word items functioning as discourse markers. Statistical analyses show that males and females in our dataset used discourse markers at a comparable rate, regardless of their personality type. Our analysis also shows that the introverted and the extroverted are not distinguished from each other in their usage of discourse markers. It also shows that the sensing individuals are distinguished from the intuiting individuals in their use of discourse markers consistently. More specifically, "uh/um" is used at the rate of 2.23 percent of the time among intuiting individuals (SD = 1.54) but is used at the rate of 1.29 percent among sensing individuals (SD = 0.81) ($p = 0.039$). Similarly, "just" is used at a higher frequency rate of 1.04 percent among intuiting individuals (SD = 0.47) than among sensing individuals who used it at the rate of 0.61 percent (SD = 0.31) ($p = 0.006$). This distinction between the sensing individuals and the intuiting individuals remains when we consider all of these markers as a whole ($p = 0.005$), in which the sensing individuals used the markers at a rate of 5.11 percent (SD = 2.26) whereas the intuiting individuals used them at a rate of 7.74 percent (SD = 2.35). We found that the two opposite groups in the thinking/feeling dimension are not distinguished from each other in their usage of discourse markers. Finally, the judging individuals and the perceiving individuals are distinguished in their use of "well". Specifically, the participants that were classified as "judging" use the discourse marker "well" 0.26 percent of the time (SD = 0.13), while those that were classified as "perceiving" use "well" 0.14 percent of the time (SD = 0.07) ($p = 0.013$).

**Table 2.** Statistical analysis results (*p*-values) on the non-lexical characteristics of the speech data.

| Gender and MBTI Personality Traits | | | Pitch (Hz) | | | Speech Rate (Words Per Minute) | | | Response Time (ms) | | | Loudness (Intensity) (dB) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Mean | SD | *p* | Mean | SD | *p* | Mean | SD | *p* | Mean | SD | *p* |
| Female (F) vs. Male (M) | F (n = 18) | | 189.6 | 21.0 | <0.001 * | 191.3 | 23.4 | 0.764 | 885 | 440 | 0.274 | 66.4 | 4.4 | 0.839 |
| | M (n = 12) | | 134.6 | 12.4 | | 193.8 | 20.8 | | 1059 | 403 | | 66.0 | 5.5 | |
| Extraversion (E) vs. Introversion (I) | E (n = 15) | F | 192.1 | 20.6 | 0.621 | 186.7 | 19.9 | 0.174 | 790 | 292 | 0.035 * | 65.9 | 5.8 | 0.683 |
| | | M | 139.0 | 14.2 | | | | | | | | | | |
| | I (n = 15) | F | 187.0 | 22.3 | 0.240 | 197.8 | 23.4 | | 1119 | 485 | | 66.6 | 3.6 | |
| | | M | 130.2 | 9.4 | | | | | | | | | | |
| Sensing (S) vs. Intuition (N) | S (n = 13) | F | 181.2 | 19.6 | 0.056 | 195.5 | 24.4 | 0.509 | 986 | 516 | 0.746 | 65.5 | 5.4 | 0.511 |
| | | M | 140.0 | 21.0 | | | | | | | | | | |
| | N (n = 17) | F | 200.0 | 18.8 | 0.621 | 189.8 | 20.6 | | 931 | 361 | | 66.8 | 4.3 | |
| | | M | 132.8 | 9.3 | | | | | | | | | | |
| Thinking (T) vs. Feeling (F) | T (n = 5) | F | 186.5 | 46.0 | 0.934 | 190.4 | 27.2 | 0.869 | 1225 | 532 | 0.224 | 65.4 | 6.8 | 0.769 |
| | | M | 132.3 | 11.1 | | | | | | | | | | |
| | F (n = 25) | F | 189.9 | 18.9 | 0.719 | 192.6 | 21.5 | | 894 | 389 | | 66.4 | 4.6 | |
| | | M | 135.3 | 13.3 | | | | | | | | | | |
| Judging (J) vs. Perceiving (P) | J (n = 18) | F | 192.0 | 16.7 | 0.618 | 186.0 | 22.7 | 0.046 * | 927 | 379 | 0.689 | 67.9 | 5.1 | 0.009 * |
| | | M | 131.0 | 9.5 | | | | | | | | | | |
| | P (n = 12) | F | 186.5 | 26.3 | 0.279 | 201.7 | 18.1 | | 996 | 506 | | 63.8 | 2.9 | |
| | | M | 141.8 | 15.8 | | | | | | | | | | |

* *p*-value < 0.05.

**Table 3.** Statistical analysis results (*p*-values) on the use of discourse markers (frequency in %).

| Gender and MBTI Personality Traits | | "Like" | | | "Kind of"/"Sort of" | | | "Uh"/"Um" | | | "Well" | | | "Just" | | | "I Mean" | | | "You Know" | | | All Combined | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Mean | SD | *p* | Mean | SD | *p* | Mean | SD | *p* | Mean | SD | *p* | Mean | SD | *p* | Mean | SD | *p* | Mean | SD | *p* | Mean | SD | *p* |
| Female vs. Male | F (n = 18) | 2.55 | 1.69 | 0.849 | 0.25 | 0.22 | 0.602 | 1.51 | 1.03 | 0.087 | 0.19 | 0.12 | 0.346 | 0.74 | 0.29 | 0.232 | 0.21 | 0.2 | 0.262 | 0.40 | 0.37 | 0.204 | 5.86 | 2.12 | 0.081 |
| | M (n = 12) | 2.67 | 1.80 | | 0.28 | 0.20 | | 2.29 | 1.65 | | 0.23 | 0.13 | | 1.03 | 0.60 | | 0.34 | 0.24 | | 0.85 | 0.78 | | 7.71 | 3.01 | |
| Extraversion vs. Introversion | Extraversion (n = 15) | 2.86 | 1.74 | 0.413 | 0.23 | 0.19 | 0.345 | 1.75 | 0.84 | 0.389 | 0.20 | 0.13 | 0.389 | 0.80 | 0.41 | 0.624 | 0.27 | 0.25 | 0.724 | 0.46 | 0.47 | 0.351 | 6.56 | 2.27 | 0.939 |
| | Introversion (n = 15) | 2.34 | 1.68 | | 0.31 | 0.23 | | 1.9 | 1.74 | | 0.23 | 0.12 | | 0.91 | 0.51 | | 0.26 | 0.19 | | 0.70 | 0.72 | | 6.64 | 3.04 | |
| Sensing vs. Intuition | Sensing (n = 13) | 2.02 | 1.78 | 0.116 | 0.28 | 0.23 | 0.967 | 1.29 | 0.81 | 0.039 * | 0.23 | 0.13 | 0.457 | 0.61 | 0.31 | 0.006 * | 0.25 | 0.21 | 0.884 | 0.43 | 0.42 | 0.403 | 5.11 | 2.26 | 0.005 * |
| | Intuition (n = 17) | 3.03 | 1.55 | | 0.27 | 0.21 | | 2.23 | 1.54 | | 0.20 | 0.12 | | 1.04 | 0.47 | | 0.27 | 0.23 | | 0.70 | 0.69 | | 7.74 | 2.35 | |
| Thinking vs. Feeling | Thinking (n = 5) | 2.46 | 1.24 | 0.806 | 0.35 | 0.29 | 0.627 | 2.15 | 1.91 | 0.957 | 0.24 | 0.10 | 0.275 | 0.77 | 0.68 | 0.327 | 0.27 | 0.23 | 0.697 | 0.88 | 0.86 | 0.243 | 7.13 | 2.83 | 0.66 |
| | Feeling (n = 25) | 2.62 | 1.80 | | 0.26 | 0.20 | | 1.76 | 1.24 | | 0.20 | 0.13 | | 0.87 | 0.42 | | 0.26 | 0.22 | | 0.52 | 0.54 | | 6.49 | 2.64 | |
| Judging vs. Perceiving | Judging (n = 18) | 2.33 | 1.20 | 0.362 | 0.30 | 0.24 | 0.819 | 1.88 | 1.2 | 0.439 | 0.26 | 0.13 | 0.013 * | 0.93 | 0.46 | 0.215 | 0.29 | 0.23 | 0.641 | 0.61 | 0.59 | 0.582 | 6.59 | 1.97 | 0.991 |
| | Perceiving (n = 12) | 3.00 | 2.26 | | 0.23 | 0.17 | | 1.73 | 1.59 | | 0.14 | 0.07 | | 0.75 | 0.45 | | 0.22 | 0.2 | | 0.54 | 0.64 | | 6.60 | 3.50 | |

* *p*-value < 0.05.

## 4. Discussion

The study observed that various speech characteristics are associated with some personality dimensions as delineated in the Myers–Briggs Type Indicator, thereby generally corroborating the relationship between speech (and, more broadly, language) and personality as noted by existing literature. Specifically, it seems that speech rate, response time, intensity, and discourse markers can be used in conjecturing a speaker's personality as informed by the MBTI categories, with each of these features associated with a more targeted personality dimension in its predictive sense.

In this section, we discuss the findings presented in the previous section in more detail. Corroborating many previous studies that observed a faster response time among extroverts [22,23,35–37], our study also found that extroverts responded 0.329 s faster than introverts in conversation ($p = 0.035$). Compared to the introverted group, extroverts are oriented more toward the outer world. This tendency arguably seems to manifest in the prompt response during an interaction, which presumably requires less inward reflection on the part of a speaker.

Our study also found that loudness is differentiated between the judging types and the perceiving types, in which the former spoke with a higher volume (approximately 4.1 dB louder) than the latter ($p = 0.009$). In addition to loudness, speech rate was also differentiated between the judging types and the perceiving types ($p = 0.046$), in which the former spoke slower than the latter. It is interesting to find that the judging types, who appear in the outside world as more structured and orderly in carrying out a variety of tasks, spoke with a higher volume and spoke slower than the perceiving types, who prefer a more adaptable and spontaneous lifestyle.

As with response time, loudness, and speech rate, discourse markers proved to be useful in reflecting different personality types. It was observed that "well", identified as "a response marker" to establish discourse coherence [38] was more frequently used by the judging types than the perceiving types ($p = 0.013$). It could be argued that the judging types—associated with conscientious types per the Big Five model [39]—are relatively more sensitive to the structure of the ongoing discourse in talk, given their tendency to organize their experiences in an orderly way. If thus posited, the judging types might arguably utilize "well" more frequently in order to achieve orderliness, or coherence, than the perceiving types, whatever this orderliness means in their minds. Furthermore, "uh/um" and "just" are more frequently used among the intuiting types than the sensing types. In general, the personality dimension of sensing/intuiting per the MBTI model seems to be the most responsive dimension in differentiating two opposite types based on the use of discourse markers. As mentioned in the previous section, when considering all the discourse markers together, the intuiting group showed a higher frequency rate than the sensing group ($p = 0.005$). This might be attributed to the characteristics of the intuiting types, who mainly are proposed to process the information not just through the sensory vehicle, but more importantly, by layering the initial information obtained via senses with interpretive meanings and patterns. These discourse markers can shape the informational or interactional structure of the discourse (e.g., indicating old information vs. new information, weakening or strengthening the effect of a statement, etc.), thereby reflecting the speaker interpretation not only on their own utterances but also on the context of the ongoing conversation. For example, we downgrade certain information in speech (e.g., "it was just a game") because we feel the need to take the emphasis away from the conveyed information. This also means that the information is conveyed in an interpretive way, with the meanings already imposed by the speaker at the time of speaking, reflecting the speaker's cognitive-psychological act of "intuiting" manifested in the ongoing talk.

Although the use of discourse markers appears in our study to be a useful predictor of the sensing/intuiting personality dimension, it remains to be seen whether this would hold when other types of words are included besides discourse markers. In addition, explaining why the use of these discourse markers is associated with sensing/intuiting dimension but not with others, such as extroversion/introversion or thinking/feeling, is

not an easy task. As speculated above, we could attempt to account for any differences in the frequency of discourse markers between the sensing group and the intuiting group by turning to the functions of these markers (i.e., approximating, filling the silence, pausing, down-grading, etc.), and propose that such functions—whichever they may be—would more likely reflect "intuiting" individuals' speech style rather than "sensing" individuals' style. For example, we might say that "intuiting" types use "just" more frequently (in other words, more likely to be down-grading), because their cognition is pre-occupied with meaning-making, or filling-in-the-lines, to a larger extent than "sensing" types. This logic, while not entirely nonsensical, is not internally supported; according to this reasoning (i.e., "just" appears more when one is cognitively "busy"), it could easily be the case that the use of "just" appears more frequently in the speech of "thinking" individuals than of "feeling" individuals, or vice versa. However, thinking types and feeling types exhibit no difference in the frequency of "like", as reported above.

At the very least, our study provides evidence in the usefulness of looking at discourse markers in personality research. Many studies have recognized the importance what people say (i.e., words) in understanding one's personality, but for the majority of these studies, the focus has been placed on words that are either semantically or functionally substantial, often leaving out the words that are "nonessential" in contributing to the overall meaning, or "habitual" phrases including "you know", "I guess", and "you see" [40] (p.46). These less-substantial words (rendered as "meaningless") are labeled as "nonfluencies" (in case of "uh" or "rr") or as "fillers" (in case of "like", "I mean", "you know", "blah", etc.) [41]. However, these words, or nonfluencies, while lacking in their substantial contribution to the meaning, occur too frequently in everyday speech to be entirely neglected. Furthermore, these discourse markers are not without functions; in the fields of sociolinguistics and discourse analysis, these words are recognized by their particular, often nonoverlapping, functions in interactional contexts. Our study corroborates previous studies in which the use of language that is not consciously controlled is found to lend insights into one's personality [42,43]. As Laserna et al. (2014) [44] show, filled pauses and discourse markers can be used to mark age groups, genders, or personality (e.g., conscientiousness) aspects. Relatedly, Iacobelli et al. (2011) [13] p. 573 noted that looking at bigrams in order to consider "words in context" yields better results in classifying personality than solely relying on thematic categories of words, thus lending us further insights into the importance of considering these discourse markers in better understanding the relationship between language and personality traits.

It should also be noted that, despite the proven effectiveness of utilizing sociolinguistic interviews in eliciting naturalistic speech, the speech data used in this study was not drawn from a completely observer-free, and therefore natural, context. Even though the participants were fairly-comfortable with the interview setting (especially attributed to the fact that the interviews were long in duration and that the participants became more and more relaxed as the interview continued), there is no way of knowing whether their interview speech was truly representative of their natural speech as would be found in an unrecorded, unobserved setting. Such environmental contexts (the interview setting, unfamiliar interviewees, etc.) could variably intervene with many aspects of the speech produced among different personality types. Perhaps certain personality types tend to be less nervous when speaking in a less natural setting, or certain personality types might be less affected by the comfort level. More studies are called for in addressing these questions. As such, the findings reported in the current study need to be validated with more participants, keeping in mind the potential impact of pseudo-naturalistic contexts in which the interviews were conducted.

Despite these limitations and reservations, our study clearly shows that language characteristics including discourse markers, speech rate, intensity, or response time warrant more of our attention in exploring various aspects of personality. The findings in this study might also be of interest for those that seek to extrapolate and utilize linguistic features in predicting personality attributes. One such effort is currently underway by the authors of this paper, in

which the aggregate of significant correlations between speech characteristics and personality dimensions is inputted into an artificial neural network, which is designed to predict personality traits only from speech data processing (the citation anonymized for review—disclosed in the cover letter). Given the increased interest in improving the quality of human-computer interactions in recent years, the current study calls for further attention to the applicability of speech data in understanding and predicting personality attributes.

The way we speak can mirror some aspects of ourselves, whether such aspects are social or psychological, including personality. While this idea resonates with us on a generally intuitive level, it is not always clear exactly what speech characteristics or linguistic features bear relevance in displaying certain aspects of personality. It is our belief that our study contributes to this line of inquiry, joining in answering the questions of the detailed relationship between language and personality.

## 5. Conclusions

The study shows that our hypothesis is confirmed for a subset of speech characteristics that we examined in their relationship with some personality traits. Specifically, we found that extroverts responded 0.329 s faster than introverts in conversation. Our study also found that the individuals that are categorized as "feeling", or those that rely on fluid contextual variables when making decisions, behave similarly to the extroverted group in terms of response time. Regardless of sex, the judging types spoke with a higher volume than the perceiving types. The personality dimension of sensing/intuiting per the MBTI model seems to be the most responsive dimension in differentiating two opposite types based on the use of discourse markers, though explaining stronger correlations between discourse markers and sensing/intuiting dimension than with other dimensions is still in question. Despite the limitations of sample size and other reservations, our study shows that language characteristics including discourse markers, pitch, intensity, or response time can be valuable indicators of various personality traits, and thus warrant more of our attention in exploring various aspects of personality.

## References

1. Addington, D.W. The relationship of selected vocal characteristics to personality perception. *Speech Monogr.* **1968**, *35*, 492–503. [CrossRef]
2. Aronovitch, C.D. The voice of personality: Stereotyped judgments and their relation to voice quality and sex of speaker. *J. Soc. Psychol.* **1976**, *99*, 207–220. [CrossRef] [PubMed]
3. Patton, M.J. A methodological preface to research on counseling. *Couns. Psychol.* **1982**, *10*, 23–26. [CrossRef]
4. Helson, R. Critics and their texts: An approach to jung's theory of cognition and personality. *J. Personal. Soc. Psychol.* **1982**, *43*, 409–418. [CrossRef]
5. Viney, L.L. The assessment of psychological states through content analysis of verbal communications. *Psychol. Bull.* **1983**, *94*, 542–563. [CrossRef]
6. Ivey, A.E. Counseling and psychotherapy: Towards a new perspective. *Couns. Psychol.* **1981**, *9*, 83–98. [CrossRef]

7.    Pennebaker, J.W.; King, L.A. Linguistic styles: Language use as an individual difference. *J. Personal. Soc. Psychol.* **1999**, *77*, 1296–1312. [CrossRef]

8.    Jurafsky, D.; Ranganath, R.; Mcfarland, D. Extracting social meaning: Identifying interactional style in spoken conversation. In Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Boulder, CO, USA, 31 May–5 June 2009.

9.    Ranganath, R.; Jurafsky, D.; McFarland, D.A. It's not You, It's Me: Detecting Flirting and Its Misperception in Speed-Dates. 2009. Available online: https://www.aclweb.org/anthology/D09-1035/ (accessed on 21 September 2021).

10.   Weninger, F.; Krajewski, J.; Batliner, A.; Schuller, B. The Voice of Leadership: Models and Performances of Automatic Analysis in Online Speeches. *IEEE Trans. Affect. Comput.* **2012**, *3*, 496–508. [CrossRef]

11.   Holtzman, N.S.; Tackman, A.M.; Carey, A.L.; Brucks, M.S.; Küfner, A.C.P.; Deters, F.G.; Mehl, M.R. Linguistic markers of grandiose narcissism: A LIWC analysis of 15 samples. *J. Lang. Soc. Psychol.* **2019**, *38*, 773–786. [CrossRef]

12.   Gawda, B. Neuroticism, extraversion, and paralinguistic expression. *Psychol. Rep.* **2007**, *100*, 721–726. [CrossRef]

13.   Iacobelli, F.; Gill, A.J.; Nowson, S.; Oberlander, J. Large Scale Personality Classification of Bloggers. In *Affective Computing and Intelligent Interaction, ACII 2011*; Lecture Notes in Computer Science; D'Mello, S., Graesser, A., Schuller, B., Martin, J.C., Eds.; Springer: Berlin, Heidelberg, 2011; Volume 6975. [CrossRef]

14.   Golbeck, J.; Robles, C.; Turner, K. Predicting Personality with Social Media. Paper Presented at CHI 2011, Vancouver, BC, Canada, 7 – 12 May 2011; Available online: http://dl.acm.org/citation.cfm?id=1979614 (accessed on 21 September 2021). [CrossRef]

15.   Beukeboom, C.J.; Tanis, M.A.; Vermeulen, I.E. The language of extraversion: Extraverted people talk more abstractly, introverts are more concrete. *J. Lang. Soc. Psychol.* **2013**, *32*, 191–201. [CrossRef]

16.   Carment, D.W.; Miles, C.G.; Cervin, V.B. Persuasiveness and persuasibility as related to intelligence and extraversion. *Br. J. Soc. Clin. Psychol.* **1965**, *4*, 1–7. [CrossRef] [PubMed]

17.   Koomen, W.; Dijkstra, W. Effects of question length on verbal behavior in a bias-reduced interview situation. *Eur. J. Soc. Psychol.* **1975**, *5*, 399–403. [CrossRef]

18.   Feldstein, S.; Sloan, B. Actual and stereotyped speech tempos of extraverts and introverts. *J. Personal.* **1984**, *52*, 188. [CrossRef]

19.   Campbell, A.; Rushton, J.P. Bodily communication and personality. *Br. J. Soc. Clin. Psychol.* **1978**, *17*, 31–36. [CrossRef]

20.   Gill, A.J.; Oberlander, J. Taking Care of the Linguistic Features of Extraversion. *Proc. Annu. Meet. Cogn. Sci. Society* **2002**.

21.   Dewaele, J.; Furnham, A. Personality and speech production: A pilot study of second language learners. *Personal. Individ. Differ.* **2000**, *28*, 355–365. [CrossRef]

22.   Siegman, A.W.; Pope, B. Effects of question specificity and anxiety-producing messages on verbal fluency in the initial interview. *J. Personal. Soc. Psychol.* **1965**, *2*, 522–530. [CrossRef]

23.   Ramsay, R.W. Speech patterns and personality. *Lang Speech* **1968**, *11*, 54–63. [CrossRef]

24.   Mairesse, F.; Walker, M.A.; Mehl, M.R.; Moore, R.K. Using linguistic cues for the automatic recognition of personality in conversation and text. *J. Artif. Intell. Res.* **2007**, *30*, 457–500. [CrossRef]

25.   Guidi, A.; Gentili, C.; Scilingo, E.P.; Vanello, N. Analysis of speech features and personality traits. *Biomed. Signal Process. Control.* **2019**, *51*, 1–7. [CrossRef]

26.   Mohammadi, G.; Vinciarelli, A. Automatic personality perception: Prediction of trait attribution based on prosodic features. *IEEE Trans. Affect. Comput.* **2012**, *3*, 273–284. [CrossRef]

27.   Polzehl, T.; Moller, S.; Metze, F. Automatically assessing personality from speech. In Proceedings of the 2010 IEEE Fourth International Conference on Semantic Computing, Pittsburgh, PA, USA, 22–24 September 2010; pp. 134–140. [CrossRef]

28.   Labov, W. *Sociolinguistic Patterns*; U. Pennsylvania Press: Oxford, UK, 1972.

29.   ELAN (Version 6.2) [Computer Software]. Nijmegen: Max Planck Institute for Psycholinguistics, The Language Archive. 2021. Available online: https://archive.mpi.nl/tla/elan (accessed on 21 September 2021).

30.   Rosenfelder, I.; Fruehwald, J.; Evanini, K.; Seyfarth, S.; Gorman, K.; Prichard, H.; Yuan, J. FAVE (Forced Alignment and Vowel Extraction) Program Suite v1.2.2 10.5281/zenodo.22281. 2014. Available online: https://www.research.ed.ac.uk/en/publications/fave-forced-alignment-and-vowel-extraction-suite-version-113 (accessed on 22 September 2021).

31.   Boersma, P.; Weenink, D. Praat: Doing Phonetics by Computer [Computer Program]. Version 6.1.39. Available online: http://www.praat.org/ (accessed on 8 February 2021).

32.   Boersma, P. Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *Proc. Inst. Phon. Sci. (Univ. Amst.)* **1993**, *17*, 97–110.

33.   Boersma, P. Acoustic Analysis. In *Research Methods in Linguistics*; Podesva, R., Sharma, D., Eds.; Cambridge University Press: Cambridge, UK, 2014.

34.   Silge, J.; Robinson, D. tidytext: Text Mining and Analysis Using Tidy Data Principles in R. *JOSS* **2016**, *1*, 37. [CrossRef]

35.   Brebner, J.; Cooper, C. The effect of a low rate of regular signals upon the reaction times of introverts and extraverts. *J. Res. Pers.* **1974**, *8*, 263–276. [CrossRef]

36.   Stahl, J.; Rammsayer, T. Extroversion-related differences in speed of premotor and motor processing as revealed by lateralized readiness potentials. *J. Mot. Behav.* **2008**, *40*, 143–154. [CrossRef]

37.   Park, J.; Lee, S.; Brotherton, K.; Um, D.; Park, J. Identification of speech characteristics to distinguish human personality of introversive and extroversive male groups. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2125. [CrossRef] [PubMed]

38.   Schiffrin, D. *Discourse Markers*; Cambridge University Press: Cambridge, UK, 1987.

39.     Furnham, A. The big five versus the big four: The relationship between the myers-briggs type indicator (MBTI) and NEO-PI five factor model of personality. *Personal. Individ. Differ.* **1996**, *21*, 303–307. [CrossRef]

40.     Seegmiller, R.A.; Epperson, D.L. Distinguishing thinking-feeling preferences through the content analysis of natural language. *J. Personal. Assess.* **1987**, *51*, 42–52. [CrossRef]

41.     Pennebaker, J.W.; Boyd, R.L.; Jordan, K.; Blackburn, K. The development and psychometric properties of LIWC2015. Manual for LIWC Program. 2015. Available online: https://repositories.lib.utexas.edu/bitstream/handle/2152/31333/LIWC2015_LanguageManual.pdf (accessed on 22 September 2021).

42.     Robinson, J.P.; Shaver, P.R.; Wrightsman, L.S. *Measures of Personality and Social Psychological Attitudes: Measures of Social Psychological Attitudes*; Academic Press: Cambridge, MA, USA, 2013.

43.     Pennebaker, J.W.; Chung, C.K.; Frazee, J.; Lavergne, G.M.; Beaver, D.I. When small words foretell academic success: The case of college admissions essays. *PLoS ONE* **2014**, *9*, e115844. [CrossRef] [PubMed]

44.     Laserna, C.M.; Seih, Y.-T.; Pennebaker, J.W. Um . . . Who like says you know: Filler word use as a function of age, gender, and personality. *J. Lang. Soc. Psychol.* **2014**, *33*, 328–338. [CrossRef]