

Identifying Personality Traits Using Overlap Dynamics in Multiparty Dialogue

Mingzhi Yu¹, Emer Gilmartin², Diane Litman¹

¹University of Pittsburgh

²Adapt Centre, Trinity College Dublin

miy39@pitt.edu, gilmare@tcd.ie, dlitman@pitt.edu

Abstract

Research on human spoken language has shown that speech plays an important role in identifying speaker personality traits. In this work, we propose an approach for identifying speaker personality traits using overlap dynamics in multiparty spoken dialogues. We first define a set of novel features representing the overlap dynamics of each speaker. We then investigate the impact of speaker personality traits on these features using ANOVA tests. We find that features of overlap dynamics significantly vary for speakers with different levels of both Extraversion and Conscientiousness. Finally, we find that classifiers using only overlap dynamics features outperform random guessing in identifying Extraversion and Agreeableness, and that the improvements are statistically significant.

Index Terms: multiparty spoken dialogue, interruption, overlap dynamics, computational paralinguistics

1. Introduction

Speech has proven to be an important key in identifying speaker personality for the psychological community. Human speech contains cues to personality traits [1] and the perception of personality based on speech is highly correlated with the perception of the whole person [2]. In recent decades, a considerable number of studies have obtained encouraging results in recognizing personality traits from human speech data [3, 4, 5, 6, 7].

Personality has been described as a constellation of traits or factors derived from the language used to describe them [8, 9]. These factors have been operationalized with different models and are frequently measured using self report quizzes such as the Big 5 (OCEAN) [10] and 16PF Questionnaire [11]. The Big 5 are five broad dimensions that can be used to describe individual differences in behavior [12]. The Big 5 traits have been described as:

- Extraversion (Extrav): energetic, emotional reactivity, assertive, sociable and talkative.
- Agreeableness (Agree): compassionate, cooperative and friendly.
- Conscientiousness (Consc): organized, dependable, self-discipline and goal-oriented.
- Neuroticism (Neuro): emotional, anxious and vulnerability.
- Openness (Open): intellectual, curious and creative.

Personality trait prediction from speech data has been performed on corpora where speakers have completed self-report personality tests at the time of recording - a process known as Automatic Personality Recognition (APR). In the absence of such self-reports, researchers have also used post hoc personality trait descriptions, made by judges reviewing audio or video

recordings, to automatically predict personality traits - a process known as Automatic Personality Perception (APP) [13]. A number of studies have used the APR [4, 14] and APP [15, 7, 5] paradigms. Mairesse et al. [4] compared the model performance of APP and APR over the same data, and found recognition of observed personality (as used in APP) was more successful than recognition of self-assessed personality (as used in APR), particularly in conversational data. They suggested that the judges observing the personality in the APP task may have used features similar to those ultimately used by the model to make their decisions, and thus were recognizing external cues interpreted as manifestations of personality rather than internal personality as measured in self-assessment. In our work, we use data from self-report Big 5 tests, and are thus recognizing internal or 'identity' personality, rather than external or 'reputation' personality.

The data used in personality trait identification has included recordings of single speakers [5, 15, 4] and spoken dialogue - both dyadic [14] and multiparty [6, 7]. Mohammadi et al. [3] and Gilpin et al. [5] used French broadcast news speech clips to perform a single speaker APP recognition experiment. The corpus contained 640 clips of 10 seconds or less representing 322 speakers, which were annotated with personality traits by a team of 11 judges. Ivanov et al. [14] used the PersIA corpus of 2 hours and 14 minutes of simulated tourist call center dialogues (119 calls and 24 identities) to perform an APR experiment. For multiparty dialogue, Pianesi et al. performed an APR experiment using the Mission Survival 2 corpus (over 6 hours) of multiparty groups performing a ranking task, comprising 12 groups of 4 participants each [6]. Valente et al. used a subset of the AMI meetings corpus [16] comprising 12 minute extracts from 32 meetings covering a total of 128 speakers [7].

The identification of personality traits from recordings has tended to use features drawn from individual speakers, either speaking alone or in groups. Lexical [4, 7], prosodic [3, 6] and acoustic features [5, 17] have been used to predict the personality trait of speakers. Valente et al. [7] included features related to interlocutor interactions when predicting personality traits in multiparty dialogue. Apart from this, few computational studies have attempted to use speaker interactions such as speech overlaps and interruptions as predictive features in identifying speaker personality traits in spoken dialogue. We investigate the effect of interlocutor traits on personality recognition with features extracted from interspeaker activity.

In the psychology literature, simple 'on-off' patterns of speech and silence and the resulting pauses, gaps and overlaps have been linked to personality traits at group and individual levels. For example, at the group level, members of less intelligent groups interrupted more frequently than members in more intelligent groups, while members of less neurotic groups interrupted more than those in more neurotic groups [18]. At the individual level, interruption rate is negatively correlated to so-

cial anxiety and speech anxiety, and positively correlated to the confidence of a speaker [19]. Individuals who appear less sociable and more assertive tend to interrupt more than those who are not [20]. Later studies also suggest that interruptions are affected by many variables including their interlocutors’ personality traits [10]. Feldstein et al. [21] found that the personality of both conversational participants influenced the frequency of onset of simultaneous speech (SS). According to him, “the extent to which an individual initiates simultaneous speech in a conversation is, indeed, influenced by aspects of his own personality. But it is also influenced ... by personality characteristics of his conversational partner.” Another interpretation of this claim is that the overall overlap dynamics of a speaker may vary depending on their interlocutors’ traits.

In this study we propose a novel method to represent the overlap dynamics of a speaker while taking account of their interlocutors’ traits. We use an APR approach to identify the Big 5 using overlap dynamics in a corpus of multiparty dialogue, containing 213 identities and 62 conversations, where each conversation is approximately 30 minutes long. To our knowledge, this is the first APR work using such overlap dynamics to predict a speaker’s Big 5 traits. Although [7] also used speaker overlap and interruption, e.g. speaker interruption frequency, our features of overlap dynamics are based on interlocutors’ traits. We use a larger corpus and more speakers than existing work using overlap information in personality trait recognition in multiparty dialogue [7]. We find overlap dynamics vary significantly for speakers with different levels of both Extraversion and Conscientiousness. Classifiers using only overlap dynamics features outperform random guessing in identifying Extraversion and Agreeableness, and the performance improvement is statistically significant.

2. Teams Corpus

The Teams Corpus¹ [22] comprises 47 hours of multiparty interaction from 62 teams (35 three-person and 27 four-person teams) playing the collaborative board game Forbidden Island. This game requires cooperation and verbal communication among the players to win as a group. The team players were 213 native speakers of American English (79M/134F) aged 18 years or older. Each player was assigned to a single team, and each team played two rounds (Game 1 and Game 2) of the game in a single session. The work presented in this paper uses only Game 1 data, where each game is approximately 30 minutes long. The corpus audio recordings were manually segmented into interpausal units (IPUs), which will serve as the basic unit of analysis in this work, using a pause length (i.e., silence) threshold of 200 milliseconds.

Participants took a pre-game survey which included a set of self-report items designed to measure the Big 5 personality traits. The 44-item Big 5 Inventory (BFI) [12] was used to score participants for each of the Big 5 traits on a scale from 1 (lowest) to 5 (highest). Figure 1 shows the score histograms for the individuals in the Teams corpus, for each trait.

We further processed the speech files to facilitate the analyses in this paper. All IPUs of less than 500 ms duration were removed, following the methodology used in [23] for work on overlap in multiparty interaction. We used Praat to create ‘floor state’ annotations to label speech, silence, and overlap activity for all participants in single labels. An alphanumeric code for each interval recorded who was speaking during the interval, or

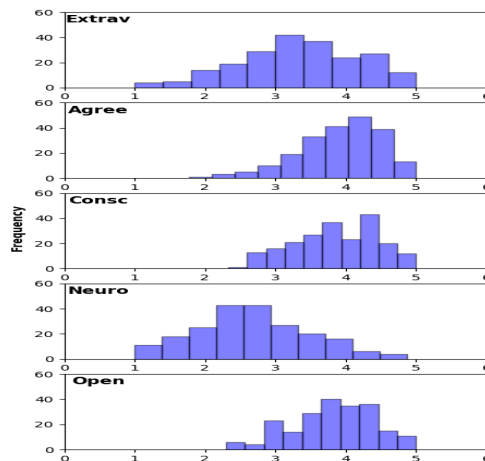


Figure 1: Big 5 score histograms for the Teams corpus.

Table 1: The number of speakers in each subgroup.

	Scores			Possessing	
	Low	Moderate	High	True	False
Extrav	56	103	54	109	104
Agree	46	133	34	111	102
Consc	40	113	60	118	95
Neuro	44	118	51	116	97
Open	47	140	26	115	98

labeled intervals of global silence (where nobody was speaking). For example, the label **aSbS** denotes that speakers **a** and **b** are speaking in overlap, while **cS** indicates that **c** is speaking alone, and **GX** denotes global silence. Any speakers not mentioned in labels are silent for the interval described. These labels could then be used with regex to generate annotations of overlap dynamics.

3. Trait Measures and Overlap Features

3.1. Personality Trait Measures

We grouped the speakers in the Teams corpus in two ways, based on their Big 5 personality trait scores. First, for a particular trait, speakers whose scores were greater than or equal to the median were considered to be those who tended to possess this trait. This grouping was used to create the overlap dynamic features described in Section 3.2. Second, for a particular trait, we automatically labeled each speaker with one of three labels: *Low* ($s < m - 0.5$), *Moderate* ($m - 0.5 \leq s \leq m + 0.5$) and *High* ($s > m + 0.5$), where s was the speaker’s score and m was the corpus median. Here we chose 0.5 as a threshold based on the previous work of [7]. These labels were used as the dependent measures for the analyses in Section 4. Table 1 shows the number of speakers in each subgroup. Note that for each trait, the group labeled with *Moderate* has the largest population based on scores. Because speakers at the median were considered to possess the trait, the number of speakers who possess a trait is slightly greater than the number who do not possess a trait.

¹<https://sites.google.com/site/teamentrainmentstudy/corpus>

3.2. Overlap Dynamics Features

We directly adopted two categories of simultaneous speech from [21] to represent overlap dynamics. For a pair of speakers, A and B, *non-interruptive simultaneous speech* (NSS) occurs when B starts speaking while A is already speaking and stops while A continues. *Interruptive simultaneous speech* (ISS) occurs when B starts speaking while A is already speaking, and continues speaking after A has stopped speaking. Figure 2 illustrates these two situations. It can be seen that there is no turn change in NSS while the turn passes from A to B in ISS.



Figure 2: *NSS and ISS illustration based on Feldstein et al. [21]. Left: non-interruptive simultaneous speech (NSS). Right: interruptive simultaneous speech (ISS).*

In [21], one of the interesting findings was that the frequency of simultaneous speech was influenced not only by the speaker’s personality characteristics, but also by the partner’s personality characteristics. For example, they found relaxed individuals were more likely to have more simultaneous speech than those who were tense. Meanwhile, if the conversational partner was talkative and attentive, it was more likely that more simultaneous speech would occur in the conversation regardless of the speaker’s personality. To our knowledge, this finding has not been previously exploited in later computational work. Thus we created a novel set of ISS and NSS features that utilize this finding by taking the personality of the conversational partner into account. For each speaker, these features represented the average frequency of ISS and NSS between this speaker and conversational partners with particular traits. For speaker i , the ISS and NSS feature of a trait (i.e., one of the Big 5), was defined in Equations 1 and 2. Here n represents the set of conversational partners that possess this trait and j represents a single speaker. ISS_{ij} and NSS_{ij} represent the frequency of ISS and NSS in the conversation. For example, *Extrav ISS* will denote the average occurrence of ISS between a speaker and speakers who possess Extraversion in the conversation. If speaker i has 5, 10, 12 ISS with speaker p , q and j correspondingly in a conversation and speaker p and q are both Extraversion, then *Extrav ISS* is $(5 + 10)/2 = 7.5$.

$$trait\ ISS = \frac{(\sum_j ISS_{ij})}{|n|}, j \in n \quad (1)$$

$$trait\ NSS = \frac{(\sum_j NSS_{ij})}{|n|}, j \in n \quad (2)$$

As speakers were randomly assigned to teams in the corpus, teams did not always include a speaker possessing a particular trait. This leads to missing *trait ISS* and *trait NSS* values for speakers who can never interact with individuals possessing specific traits. Our handling of missing feature values is discussed in Section 4.

The ISS and NSS features capture the overlap dynamics of a speaker based on dyadic aspects of team interactions. In multi-party conversation, simultaneous speech also occurs across more than 2 speakers. Therefore, we also created features to represent overlap involving more than two speakers.

Table 2: *One-way ANOVA significant results. * if $p < 0.05$ and ** if $p < 0.01$. M: Moderate, H: High, L: Low. Sample size shows the total number of samples used in ANOVA after eliminating missing data.*

	Significant Features	Post hoc	Mean			Sample Size
			L	M	H	
Extrav	2 spks overlap**	M < H, L < H	182.4	197.2	243.2	213
	3 spks overlap*	M < H	48.8	47.3	65.0	213
	Extrav ISS*	L < H, M < H	65.1	67.0	84.6	180
	Consc ISS*	L < H	59.0	66.9	78.8	181
	Neuro ISS*	M < H	63.4	64.7	80.3	182
	Open ISS*	M < H	60.4	60.9	77.8	182
	Open NSS*	L < H, M < H	23.4	23.2	30.5	182
Consc	Agree NSS*	M < L	29.3	21.7	25.9	174

Here we were particularly interested in overlaps involving 2 speakers and 3 or more speakers since our corpus consisted of 3-party and 4-party conversations. We amalgamated 3 and 4 speaker overlap as 4-speaker overlap was extremely rare in the data – accounting for less than 0.2% of total overlap in the corpus. *2 spks overlap* denotes the number of overlaps involving 2 speakers (including the speaker). *3+ spks overlap* denotes the number of overlaps involving at least 3 speakers (including the speaker).

4. Data Analysis and Model Building

4.1. Differences in Overlap Dynamics by Trait Strength

To investigate whether the values of our overlap dynamics features differ significantly across subjects in the *Low*, *Moderate* and *High* trait groups introduced in Section 3.1, we performed a one-way ANOVA test for each of the Big 5 traits. For each analysis, we eliminated missing values. The significant results are shown in Table 2. Overlap dynamics were found to significantly vary with Extraversion and Conscientiousness for *Low*, *Moderate* and *High*. For Extraversion, *High* tends to be involved in more 2 speaker and 3+ speaker overlaps than other labeled groups. Also, *High* of Extraversion generally tends to have more ISS with other traits and NSS with Openness. For Conscientiousness, *Low* tends to have more NSS with Agreeableness. Overall, Extraversion emerges as the trait that shows a diverse pattern in different levels. Extroverts overlap with other speakers more frequently than introverts in the conversation. Except for Agreeableness, the interruptive activities occurred frequently between extroverts and other speakers. Compared to introverts, extroverts generally tend to overlap with other speakers regardless of their personalities. These findings have matched some characteristics of extraversion such as being talkative and energetic, which were mentioned in related literature [9]. Being low in Conscientiousness can be characterized as being careless and inefficient. The ANOVA result suggests that careless individuals tend to have more non-interruptive overlap with friendly and cooperative persons compared to those who are organized and self-disciplined.

4.2. Predicting Personality Traits using Overlap Dynamics

Moving from ANOVA analyses to prediction, we constructed a separate Naive Bayes classifier² for each of the Big 5, to predict one of the following trait labels from the overlap dynamics features: *Low*, *Moderate* and *High*. To handle samples with missing feature values, we evolved the imputation strategy in data splitting with K nearest neighbors. We took three steps to

²We also experimented with SVM and Decision Trees. Here we report Naive Bayes as it achieved the best performance using our data.

Table 3: Predicting Low, Moderate and High. Evaluation metrics: macro averaged Precision (P), Recall (R) and F1. * if $p < 0.05$, ** if $p < 0.01$. Improved performance compared to random is shown in boldface.

	Naive Bayes			Random Guessing		
	P	R	F1	P	R	F1
Extrav	0.32	0.38**	0.31	0.31	0.30	0.30
Agree	0.37	0.38*	0.36**	0.32	0.31	0.28
Consc	0.37	0.32	0.30	0.34	0.33	0.31
Neuro	0.19	0.24	0.21	0.34	0.33	0.32
Open	0.26	0.35	0.29	0.33	0.31	0.28

partition the dataset into 70% for training and 30% for testing. We first divided the data set into two subsets, A (containing only samples with no missing values) and B (the rest of the samples). We randomly drew from subset A a number of samples corresponding to 30% of the size of the entire dataset. This formed our test set. We then added the remainder of subset A to subset B to form the training set.

We imputed the training set using k-Nearest Neighbour (KNN). To alleviate the class imbalance caused by sampling from different populations, we created 10 train/test splits with the same techniques. The average performance of classifiers over the 10 splits was reported in the evaluation. Also, the performance was compared with a baseline classifier that predicts labels randomly.

The evaluation results are shown in Table 3. We compared the performance of the classifier and the baseline over the 10 splits by t-test. The classifier for Agreeableness outperformed the baseline, and the improvements in recall and F1 are statistically significant. The Extraversion classifier also outperformed the baseline but only Recall showed significant improvement. Precision for Conscientiousness, and Recall and F1 for Openness were slightly improved but not significantly. Overall, the performance of the classifiers did not appear to be adequately strong for all the Big 5.

We noticed the evaluation of *Low* and *High* showed promising performance depending on the traits. For example, for *High* in Extraversion, Precision (0.24), Recall (0.47) and F1 (0.31) were all statistically significantly higher than Precision (0.14), Recall (0.24) and F1 (0.17) for the baseline. To further investigate how the features predict *Low* and *High*, we experimented with only these two labels, following [7], and saw promising prediction accuracy. After removing all *Moderate* samples, we trained and tested Naive Bayes classifiers. The results are shown in Table 4. The absolute performance of the classifiers and the baseline were both improved compared to the prediction of *Low*, *Moderate* and *High*. This finding implies that it is easier for classifiers to distinguish between just *Low* and *High*. With respect to the prediction outcome after utilizing overlap dynamics features, we observed a performance improvement in predicting *Low* and *High* of Extraversion compared to the baseline. The precision, recall and F1 are all improved, and the improvement of F1 is statistically significant. In distinguishing between *Low* and *High* of Agreeableness, the classifier outperformed the baseline significantly. Note that there were some improvements in predicting *Low* and *High* of Conscientiousness and Open but they were not statistically significant.

In conclusion, only using the overlap dynamics features, Naive Bayes classifiers can achieve up to 29% relative improvement in F1 compared to the baseline when identifying three labels for a trait (with the highest relative improvement for Agreeableness). The largest relative F1 improvement in identifying

Table 4: Predicting Low and High. Evaluation metrics: macro averaged Precision (P), Recall (R) and F1. * if $p < 0.05$, ** if $p < 0.01$. Improved performance compared to random is shown in boldface.

	Naive Bayes			Random Guessing		
	P	R	F1	P	R	F1
Extrav	0.57	0.57	0.56*	0.51	0.51	0.49
Agree	0.62*	0.63*	0.58*	0.54	0.54	0.53
Consc	0.60	0.53	0.51	0.56	0.56	0.55
Neuro	0.31	0.36	0.29	0.45	0.45	0.44
Open	0.49	0.49	0.47	0.49	0.49	0.46

just the low and high levels is 14% for Extraversion, compared to the baseline.

5. Discussion and Future Work

In this work, we presented an APR approach that uses overlap dynamics in multiparty dialogue. The APR experiments were conducted over a multiparty spoken corpus that contains 213 identities and 62 conversations, where each conversation is approximately 30 minutes long. We introduce a novel representation of overlap dynamics that takes the partner’s trait into account. The ANOVA result reveals that the interruption and overlap behaviors may vary between different levels of personality traits. Overall, we found overlap dynamics can be used as indicators of some traits. With respect to APR, models utilizing overlap dynamics features show promising performance. The predictions of Agreeableness and Extraversion show the most significant improvement. Note that in the earlier APP work [7], the recognition of Agreeableness also shows statistically significant improvement when using their participant interaction features, which only reflect the speaker’s overlapping activities but not the overlapping interaction between partners with different traits. It is unclear if this similar finding is a coincidence or caused by the underlying correlation between Agreeableness and a speaker’s overlapping behavior. Future work is needed.

The current method to determine the possession of traits relies on the median of scores. Since the speaker is also included in the population, it may impact the overall population distribution and result in a potential leak of information in the processing. In the future, we intend to improve the independence of feature extraction. We plan to apply techniques of data splitting such as splitting speakers and their conversational partners. Meanwhile, the current feature extraction of ISS and NSS is based on a supervised dataset, which contains the knowledge of all speakers’ personality, including the personality of their conversational partners. In the future, we are interested in approaching the same problem in an unsupervised manner without using the traits of partners. The presented study mainly shows that overlap dynamics provided some signal of speaker personality. To further study overlap dynamics, we intend to examine its effectiveness in conjunction with other well-established features in the literature.

6. Acknowledgements

This work is supported by the National Science Foundation under Award # 1420784. Emer Gilmartin is supported by the ADAPT Centre for Digital Content Technology, which is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

7. References

- [1] K. R. Scherer and H. Giles, *Social markers in speech*. Cambridge University Press, 1979, vol. 6.
- [2] P. Ekman, W. V. Friesen, M. O'sullivan, and K. Scherer, "Relative importance of face, body, and speech in judgments of personality and affect." *Journal of personality and social psychology*, vol. 38, no. 2, p. 270, 1980.
- [3] G. Mohammadi and A. Vinciarelli, "Automatic personality perception: Prediction of trait attribution based on prosodic features," *IEEE Transactions on Affective Computing*, vol. 3, no. 3, pp. 273–284, 2012.
- [4] F. Mairesse, M. A. Walker, M. R. Mehl, and R. K. Moore, "Using linguistic cues for the automatic recognition of personality in conversation and text," *Journal of artificial intelligence research*, vol. 30, pp. 457–500, 2007.
- [5] L. H. Gilpin, D. M. Olson, and T. Alrashed, "Perception of speaker personality traits using speech signals," in *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, 2018, p. LBW514.
- [6] F. Pianesi, N. Mana, A. Cappelletti, B. Lepri, and M. Zancanaro, "Multimodal recognition of personality traits in social interactions," in *Proceedings of the 10th international conference on Multimodal interfaces*. ACM, 2008, pp. 53–60.
- [7] F. Valente, S. Kim, and P. Motlicek, "Annotation and recognition of personality traits in spoken conversations from the ami meetings corpus," in *Thirteenth Annual Conference of the International Speech Communication Association*, 2012.
- [8] R. B. Cattell, "The description of personality: Basic traits resolved into clusters." *The journal of abnormal and social psychology*, vol. 38, no. 4, p. 476, 1943.
- [9] W. T. Norman, "2800 personality trait descriptors—normative operating characteristics for a university population." 1967.
- [10] J. Goldberg, "Interrupting the discourse on interruptions: An analysis in terms of relationally neutral, power- and rapport-oriented acts," *Journal of Pragmatics*, vol. 14, pp. 883–903, 12 1990.
- [11] H. E. Cattell and A. D. Mead, "The sixteen personality factor questionnaire (16pf)," *The SAGE handbook of personality theory and assessment*, vol. 2, pp. 135–178, 2008.
- [12] O. P. John, E. Donahue, and R. Kentle, "The big five inventory: Versions 4a and 54 [technical report]," *Berkeley, CA: University of California, Institute of Personality and Social Research*, 1991.
- [13] B. Schuller, S. Steidl, A. Batliner, E. Nöth, A. Vinciarelli, F. Burkhardt, R. van Son, F. Weninger, F. Eyben, T. Bocklet *et al.*, "A survey on perceived speaker traits: Personality, likability, pathology, and the first challenge," *Computer speech & language*, vol. 29, no. 1, pp. 100–131, 2015.
- [14] A. V. Ivanov, G. Riccardi, A. J. Sporka, and J. Franc, "Recognition of personality traits from human spoken conversations," in *Twelfth Annual Conference of the International Speech Communication Association*, 2011.
- [15] G. Mohammadi, A. Vinciarelli, and M. Mortillaro, "The voice of personality: Mapping nonverbal vocal behavior into trait attributions," in *Proceedings of the 2nd international workshop on Social signal processing*. ACM, 2010, pp. 17–20.
- [16] J. Carletta, S. Ashby, S. Bourban, M. Flynn, M. Guillemot, T. Hain, J. Kadlec, V. Karaiskos, W. Kraaij, M. Kronenthal *et al.*, "The ami meeting corpus: A pre-announcement," in *International Workshop on Machine Learning for Multimodal Interaction*. Springer, 2005, pp. 28–39.
- [17] F. Eyben, M. Wöllmer, and B. Schuller, "Opensmile: the munich versatile and fast open-source audio feature extractor," in *Proceedings of the 18th ACM international conference on Multimedia*. ACM, 2010, pp. 1459–1462.
- [18] Y. Rim, "Personality variables and interruptions in small discussions." *European Journal of Social Psychology*, 1977.
- [19] M. Natale, E. Entin, and J. Jaffe, "Vocal interruptions in dyadic communication as a function of speech and social anxiety." *Journal of Personality and Social Psychology*, vol. 37, no. 6, p. 865, 1979.
- [20] L. F. Robinson and H. T. Reis, "The effects of interruption, gender, and status on interpersonal perceptions," *Journal of nonverbal behavior*, vol. 13, no. 3, pp. 141–153, 1989.
- [21] S. Feldstein *et al.*, "Personality and simultaneous speech." 1974.
- [22] D. Litman, S. Paletz, Z. Rahimi, S. Allegretti, and C. Rice, "The teams corpus and entrainment in multi-party spoken dialogues," in *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 2016, pp. 1421–1431.
- [23] E. Kurtić, G. J. Brown, and B. Wells, "Resources for turn competition in overlap in multi-party conversations: speech rate, pausing and duration," in *Eleventh annual conference of the international speech communication association*, 2010.