Acoustic-Prosodic Automatic Personality Trait Assessment for Adults and Children

Rubén Solera-Ureña[1]*, Helena Moniz[1,2], Fernando Batista[1,3], Ramón Fernández-Astudillo[1,4], Joana Campos[5,6], Ana Paiva[5,6], and Isabel Trancoso[1,6]

[1] Spoken Language Systems Laboratory, INESC-ID Lisboa, Portugal
{rsolera, helenam, fmmb, ramon.astudillo, imt}@l2f.inesc-id.pt
[2] FLUL / CLUL, Universidade de Lisboa, Portugal
[3] ISCTE-IUL – Instituto Universitário de Lisboa, Portugal
[4] Unbabel Inc., Lisboa, Portugal
[5] Intelligent Agents and Synthetic Characters Group, INESC-ID Lisboa, Portugal
joana.campos@gaips.inesc-id.pt, ana.paiva@inesc-id.pt
[6] Instituto Superior Técnico, Universidade de Lisboa, Portugal

Abstract. This paper investigates the use of heterogeneous speech corpora for automatic assessment of personality traits in terms of the *Big-Five* OCEAN dimensions. The motivation for this work is twofold: the need to develop methods to overcome the lack of children's speech corpora, particularly severe when targeting personality traits, and the interest on cross-age comparisons of acoustic-prosodic features to build robust paralinguistic detectors. For this purpose, we devise an experimental setup with age mismatch utilizing the Interspeech 2012 Personality Sub-challenge, containing adult speech, as training data. As test data, we use a corpus of children's European Portuguese speech. We investigate various features sets such as the Sub-challenge baseline features, the recently introduced eGeMAPS features and our own knowledge-based features. The preliminary results bring insights into cross-age and -language detection of personality traits in spontaneous speech, pointing out to a stable set of acoustic-prosodic features for Extraversion and Agreeableness in both adult and child speech.

Keywords: Computational Paralinguistics, Automatic Personality Assessment, OCEAN, Cross-lingual, Cross-age

1 Introduction

The intents, emotions, and even personality traits of a speaker are coded in paralinguistic information, beyond the linguistic structures of a language. The **OCEAN** (*Big-Five*) dimensions of personality traits are a psychological construct summarized as follows: Openness (artistic, imaginative, original), Conscientiousness (organized, efficient, thorough), Extraversion (energetic, outgoing, talkative), Agreeableness (kind, generous, sympathetic), and Neuroticism (anxious, self-pitying, worrying).

The analysis of personality traits has a plethora of applications, e.g., discriminating natural from disordered behaviors or automatically assessing personality traits, either in human-human communications or in human-computer interactions. Much of the literature on automatic processing of personality traits is still mostly focused on assessing and detecting the traits based on several sets of distinct features. Artificial intelligence applications are, however, giving steps towards displaying robots and virtual agents with certain traits to better interact with humans, making the communication more idiosyncratic and tuned to the paralinguistic fingerprints of an interlocutor.

The automatic assessment/detection of personality traits is still a very challenging task, either due to the individual spectrum of a speaker, or to the spectrum of the trait itself: whenever the richness of a person is defined in 5 classes, it may not cover all the sub-specifications or the boundaries between such classes, as psychological studies have been pointing out [4]. It is clear in the literature that some traits can be more easily recognized by means of automatic procedures than others, but this fact may vary according to the data and the methodologies applied (see [21] for a survey). Moreover, it has been timidly pointed out that different personality dimensions/traits are revealed in spontaneous speech by means of different sets of representative acoustic/prosodic features [11, 14–16, 21], but exhaustive categorizations of such features and studies on their impact across ages, cultures, etc. are still missing.

Psychological studies have shown a strong debate between change and continuity of personality traits from childhood to adult age or even elderly in longitudinal studies [3, 7, 9, 18]. The studies in [3] show that children's personality traits are linked to that ones displayed in adult age; for instance: *"When observed at age 3, children classified as Inhibited were shy, fearful, and socially ill at ease. At age 26, they were characterized by an overcontrolled and nonassertive personality style; they expressed little desire to exert influence over others and reported taking little pleasure in life"*.

Reasoning about these findings from an automatic processing point of view, we could expect to find similar sets of acoustic/prosodic features for a given trait in adult and children speech, which could lead to reasonable performance rates for the classification of personality traits across languages and ages. This is the main motivation for our work here that targets the experimental evaluation of personality models trained on French adults' speech on a completely different corpus consisting of Portuguese children's speech.

2 Cross-Age and Cross-Language Datasets

Two very different speech corpora have been used in this work, namely the Speaker Personality Corpus (SPC) [13, 19, 20] and the Game-of-Nines (GoN) corpus [2]. The markedly different characteristics of the speakers in these two corpora (adults vs. children, French vs. Portuguese speakers, respectively) constitute a good basis to investigate whether we can leverage the presence of common, language- and age-independent acoustic and prosodic cues to detect/assess basic personality traits across heterogeneous groups of target speakers. The more populated SPC database was used in this work to learn (and evaluate) statistical models for the binary classification tasks corresponding to each personality trait in the *Big-Five* model (**OCEAN**). These models were then used to automatically assess the perceived personality traits of the children present in the GoN corpus. Below, we present a description of these two speech data sets.

2.1 Speaker Personality Corpus

The Speaker Personality Corpus consists of 640 speech files from 322 different adult individuals, collected from the French news bulletin of Radio Suisse Romande, Switzerland. Each file contains 10 seconds of speech from just one speaker (around 1 hour and 40 minutes in total). All the files were independently assessed by 11 judges (non-French speakers) using the BFI-10 personality questionnaire [17]. For each file, final labels for the *Big-Five* dimensions are calculated by a majority vote procedure: for each trait, a high/low level (O/NO, C/NC, E/NE, A/NA, N/NN) is assigned if at least 6 judges scored it above/below their personal averages for that trait.

The SPC corpus was used in the Interspeech 2012 Speaker Trait Challenge-Personality Sub-challenge [19, 20]. On that occasion, the corpus was split into 3 speaker-independent train, development and test subsets consisting of 256, 183 and 201 files, respectively. In this work, the same experimental setup has been adopted.

2.2 Game-of-Nines Corpus

The Game-of-Nines corpus was originally designed to study how conflict unfolds in social interactions by looking at behavioral cues (e.g. gaze) in a mixed-motive social interaction (i.e. a scenario with competitive and cooperative incentives) with children. It comprises synchronized video- and audio-recordings of 11 dyadic sessions with 22 Portuguese children (13 girls and 9 boys) aged 10 to 12 years-old playing a competitive-cooperative bargaining card game (a modified version of the *Game of Nines* [8]). The duration of the recordings vary between 9 and 18.6 minutes, with an average duration of 12.8 minutes and a total of 2 hours and 20 minutes. Personality annotations for the children that participated in the experiments were provided with the database; the self-administered PBPS-C v2 personality questionnaire for children [10] was employed in that case. However, the original annotations were discarded in this work and an alternative annotation process was conducted. The reason behind this is that the original personality scores, directly derived from children's self-reports, might be biased towards the ideal personal image that the children would like to project. Finally, manual transcriptions of the conversations are also provided with this database.

The original Game-of-Nines database was pre-processed in order to adapt it for our purposes. Firstly, all the video information was discarded for this work and just the audio tracks were used. Secondly, the speech transcriptions were used to identify and extract all the speech segments corresponding to each child.

Those parts with overlapped conversations (both playmates speaking at the same time) were removed in order to avoid processing mixed speech segments. After this pre-processing procedure, two different speech subsets were generated:

1. *GoN-complete*: for each child, all their speech segments in the game session were concatenated together in one single speech file. As a results, the *GoN-complete* subset consists of 22 files ranging from 49 seconds to 8.1 minutes of speech (average duration of 4.2 minutes).
2. *GoN-20seconds*: for each child, 4 different files with around 20 seconds of speech each were generated by concatenating their longer speech segments in the game. Very short segments (below 2 seconds) were discarded in order to avoid an excessive variability in the speech characteristics. Also, a speaker-balanced subset was considered advisable. With these restrictions, just 4 files could be generated for the majority of the children, while just 2 files could be generated for one of the participants. As a result, the *GoN-20seconds* subset consists of 86 files with an approximate duration of 20 seconds.

Finally, both speech subsets were independently annotated in terms of the *Big-Five* personality dimensions by a professional psychologist using the BFI-10 personality questionnaire. These annotations have been used as the ground-truth labels in this work. The psychologist that participated in this work has substantial experience in annotating data and has already participated in several research projects in this field.

These two GoN subsets have been used in this work to study how personality models built up from French adults' speech can be used to assess the *Big-Five* dimensions of personality of Portuguese children. Additionally, differences in the average length of the files in these two subsets allow a comparison of the effect of short versus long acoustic cues on the personality assessment systems.

3 Classification System for Paralinguistics

This section presents a description of the speech representations and statistical models used in this work.

3.1 Recent Progress in Parameterizations for Paralinguistics

In our experiments, we used two sets of features extracted with openSMILE [6], and a set of knowledge-oriented features known in the literature to have impact on the personality classification tasks tackled here, henceforth referred to as knowledge-based features. The Interspeech 2012 Speaker Trait Challenge-Personality Sub-challenge feature set consists of 6125 features, and has been used in the present work to provide a set of baseline results (IS2012). We have also used the eGeMAPS feature set [5] (an extended version of GeMAPS - Geneva Minimalistic set of Acoustic Parameters for Voice Research and Affective Computing) that consists of 88 features, well-known for their usefulness in a wide range of paralinguistic tasks.

Our knowledge-based features (KB-features) are based on a phone tokenization of the speech files using the neural network-based acoustic models of the AUDIMUS speech recognizer [12]. The phonetic tokenizations provide phone alignments for each speech file, which can be used to extract duration-related features and to generate more advanced features. In this way, for instance, it is possible to extract the silence ratio, speech duration ratio, and speech rate features in terms of phones per second. The phone tokenizations also provide us with means to characterize each speech segment using n-grams of phones. Based on these tokenizations, we then derive *Inter Pausal Units* (IPUs), that consist of sequences of phones delimited by silences.

The experiments presented in this work use a set of 40 knowledge-based features, including duration of speech with and without internal silences, and tempo measurements such as speech and articulation rates (number of phones or syllables divided by the duration of speech with and without internal silences, respectively) and phonation ratio (duration of speech without internal silences divided by the duration of speech including internal silences). Other features involve pitch (f0), energy, jitter and shimmer, including pitch and energy average, median, standard deviation, dynamics, range, and slopes, both within and between IPUs [1]. Pitch related features were calculated based on semitones rather than frequency. On top of such features, we extracted elaborated prosodic features for the whole sentence involving the sequence of derived IPUs, that were expressed in terms of standard deviation and slope. The Snack Sound Toolkit[1] was used to extract the pitch and energy from the speech signal. Jitter and shimmer were extracted from openSMILE low-level descriptors. For the time being KB-features are still not extensive and must be used in combination with other features in order to achieve improved performances.

3.2 Models for Personality Assessment

Speech corpora annotated in terms of personality traits are scarce and generally small. Therefore, the application of too elaborated machine learning methodologies to learn complex models is neither advisable nor viable in many cases. In general, complex models are described by a high number of parameters that require larger training data sets; otherwise, overfitting may easily occur.

In this work, the same experimental setup as that employed in the Interspeech 2012 Speaker Trait Challenge-Personality Sub-challenge [19, 20] has been adopted. The models used in this work are linear support vector machines (SVM) trained by means of the well-known SMO algorithm. Logistic functions have been fitted to the SVM soft outputs in order to transform them into pseudo-posterior probabilities. Two different types of feature normalization ($[0, 1]$ range, and zero-mean and unit-variance) have been used.

Each trait in the *Big-Five* personality model has been considered as an independent binary classification problem, where the goal is to assign (classify) a high/low level on that trait (O/NO, C/NC, E/NE, A/NA, N/NN) to every

[1] http://www.speech.kth.se/snack/

speech file. Thus, five different models were trained in this work corresponding to Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. The SVM models were trained using data from the SPC corpus. Firstly, a grid-search process was applied to find the optimal value for the complexity parameter C of the SVMs. For this purpose, different SVMs were trained on the training subset (values for C in the range 10^{-5} to 10 were used) and the associated unweighted average recall (UAR) on the development subset were calculated. The value for C providing the higher UAR on the development subset was then selected as the optimal value. Then, the training and development subsets were merged together and the definitive SVM models were trained on this data set, using the selected values for C. Finally, the UAR on three different test sets (SPC test set, *GoN-complete* and *GoN-20seconds* sets) is calculated to evaluate the models on both same- and cross-language conditions.

4 Experiments and Results

This section is devoted to the presentation and discussion of the experimental results obtained by the systems described previously. Three key aspects have been studied in this work: 1) the performance of three different parameterizations; 2) the consequences of using cross-lingual and cross-age speech corpora for the automatic assessment of personality traits; and 3) the effects of the use of short versus long acoustic-prosodic cues on this task. The results are presented in terms of unweighted average recall (UAR) and accuracy (Acc). The UAR is a fairer measure (and thus preferred) when the data present a substantial imbalance, which is the case of the data sets employed here. The number of examples in each class (high or low level for a trait) for the SPC corpus and for the two GoN subsets are shown in Tables 1 and 2, respectively.

Table 1. Number of examples in each class (high/low level for a trait) for the SPC corpus.

Trait	SPC Train					SPC Devel.					SPC Test				
	O	C	E	A	N	O	C	E	A	N	O	C	E	A	N
#High	97	110	121	139	140	70	81	92	79	88	80	99	107	105	90
#Low	159	146	135	117	116	113	102	91	104	95	121	102	94	96	111

Table 3 shows the results obtained on the SPC corpus used in the Interspeech 2012 Speaker Trait Challenge-Personality Sub-challenge. We present here development and test results for the IS2012 baseline features (6125), eGeMAPS features (88), and the combined eGeMAPS+KB-based features (128). Also, the optimal values for the complexity parameter C of the SVMs, selected from the results on the development set, are presented.

We can see that the baseline IS2012 feature set achieves the best test results (in terms of UAR) for Conscientiousness and Agreeableness, while the eGeMAPS

Table 2. Number of examples in each class (high/low level for a trait) for the *GoN-20seconds* and *Gon-complete* subsets.

	GoN-20seconds					*GoN-complete*				
Trait	O	C	E	A	N	O	C	E	A	N
#High	9	52	53	51	33	6	14	18	13	13
#Low	77	34	33	35	53	16	8	4	9	9

Table 3. Results achieved on the SPC data set.

	IS2012					eGeMAPS					eGeMAPS+KB-features				
	C	Devel.		Test		C	Devel.		Test		C	Devel.		Test	
Trait		UAR	Acc	UAR	Acc		UAR	Acc	UAR	Acc		UAR	Acc	UAR	Acc
O	1E-05	63.5	66.7	58.6	60.2	3E-04	66.4	68.9	59.3	62.2	1E-03	62.1	63.9	53.9	55.7
C	1E-02	74.5	74.9	80.1	80.1	1E+00	74.8	74.9	76.1	76.1	1E-01	73.9	73.8	79.1	79.1
E	1E-04	82.0	82.0	75.4	75.6	3E-03	83.0	83.1	72.0	72.1	1E-03	83.0	83.1	77.2	77.6
A	3E-05	68.1	66.1	62.5	62.7	3E-02	65.8	63.4	59.4	59.7	1E-04	67.4	65.0	57.8	58.2
N	1E-04	69.1	69.4	63.4	63.7	1E-05	69.9	69.9	63.7	64.2	1E-01	70.9	71.0	63.4	63.7

features achieve best results for Openness and Neuroticism, and the combined eGeMAPS+KB-based feature set provides the best result for Extraversion. It is worth mentioning the poor results of the eGeMAPS features for Conscientiousness, Extraversion and Agreeableness, which suggest that such a small feature set is not able to fully capture the essence of those personality traits. For Conscientiousness and Extraversion, the addition of the KB-based features leads to a considerable improvement in the UAR. This shows that the particular nature of the KB-based features, based on an initial phone tokenization of the speech files, endows them with the capability to complement the eGeMAPS feature set in certain cases. However, the addition of the knowledge-based features has the opposite effect for Openness and, to a lesser extent, Agreeableness.

From these results, we must emphasize that the task of automatic personality assessment based on speech information requires a careful selection of acoustic and prosodic feature sets specific for each personality trait. Furthermore, it is shown that a wise selection of a modest number of adequate features based on task-specific knowledge can lead to similar or even better results that those obtained with huge, generic feature sets.

Tables 4 and 5 show the results obtained on the *GoN-20seconds* (86 speech files) and *GoN-complete* (22 files) subsets, respectively. The personality models employed in these experiments are those previously trained on the SPC corpus. To the best of our knowledge, our approach is quite novel and it does not exist any previous work with a similar cross-age and cross-language setup that could be used as a proper baseline for the results presented here. It is also worth mentioning that the results for the Openness trait should be analyzed with certain reservations, since it was very difficult for the psychologist involved in this work to assess the items in the BFI-10 personality questionnaire related to the Openness trait (namely, "The speaker has few artistic interests" and

Table 4. Results achieved on the *GoN-20seconds* data set.

Trait	IS2012			eGeMAPS			eGeMAPS+KB-features		
	C	Test		C	Test		C	Test	
		UAR	Acc		UAR	Acc		UAR	Acc
O	1E-05	50.0	89.5	3E-04	49.4	88.4	1E-03	50.0	89.5
C	1E-02	50.8	43.0	1E+00	50.0	50.0	1E-01	44.9	40.7
E	1E-04	64.5	60.5	3E-03	57.9	55.8	1E-03	61.1	57.0
A	3E-05	65.0	68.6	3E-02	61.5	64.0	1E-04	61.7	66.3
N	1E-04	50.0	38.4	1E-05	48.6	50.0	1E-01	50.3	46.5

Table 5. Results achieved on the *GoN-complete* data set.

Trait	IS2012			eGeMAPS			eGeMAPS+KB-features		
	C	Test		C	Test		C	Test	
		UAR	Acc		UAR	Acc		UAR	Acc
O	1E-05	80.2	86.4	3E-04	60.4	72.7	1E-03	50.0	72.7
C	1E-02	50.0	63.6	1E+00	64.3	54.5	1E-01	48.2	40.9
E	1E-04	66.7	45.5	3E-03	59.7	50.0	1E-03	69.4	50.0
A	3E-05	51.3	54.5	3E-02	73.5	72.7	1E-04	67.9	68.2
N	1E-04	50.0	59.1	1E-05	33.8	31.8	1E-01	46.2	54.5

"The speaker has an active imagination"). In most of the cases, the psychologist assigned default values to these items, which harms the statistical relevance of the results for Openness.

The most relevant outcome from these experiments is that reasonable and consistent results across different setups (three different feature sets and short vs. long acoustic-prosodic cues) are obtained for Extraversion and Agreeableness. The UAR values for these personality traits are consistently above 60% in most of the cases, with a maximum value of 73.5% for Agreeableness on the *GoN-complete* subset, using the eGeMAPS features. These results point out to the existence of a stable set of acoustic-prosodic features for these traits in both adult and children speech, which supports our aim of using cross-language and cross-age speech corpora for the assessment of personality traits in those situations where data are specially scarce. On the other hand, the results do not show relevant performances for Conscientiousness and Neuroticism. In our opinion, the reasonable and promising results obtained for Extraversion and Agreeableness are strongly linked to the specific design and characteristics of the Game-of-Nines study, where different classmate dyads play a competitive-cooperative bargaining card game for a final reward. This particular situation might favor more noticeable expressions of those traits more closely related to this specific dyadic interaction setup, such as Extraversion and Agreeableness.

Tables 4 and 5 do not show clearly better results for any of the three parameterizations. If we focus on Extraversion and Agreeableness, the IS2012 feature set achieves better results on the *GoN-20seconds* subset, while the eGeMAPS and combined eGeMAPS+KB-based feature sets perform better on the *GoN-complete* subset. Finally, the experimental results show that, in general, better results are achieved on the *GoN-complete* subset for those traits. This result points out the limitations of our 20 seconds-long speech segments constructed by concatenating several smaller segments.

5 Conclusions and Future Work

This work investigates the use of heterogeneous speech corpora for automatic personality assessment tasks. The main motivation for this work is the need to develop methods to overcome the lack of children's speech corpora in this field. With this purpose, we evaluated the use of personality models trained on French adults' speech to classify a completely different corpus of Portuguese children's speech. Our preliminary results bring insights into cross-age and cross-language detection of personality traits in spontaneous speech. The reasonable performance rates obtained for Extraversion and Agreeableness point out to a stable set of acoustic-prosodic features in both adult and children speech. Also, the importance of a sensible selection of specific feature sets for each personality trait is shown in this paper.

Further work will be carried out in this research line. In particular, the acquisition of more speech data, together with the inclusion of more assessors to improve the statistical significance of our personality annotations, is of paramount importance to perform more exhaustive and relevant experimentation. This will allow the study of more elaborated feature selection procedures and learning methodologies with the aim of developing robust personality assessment systems.

Acknowledgments. Work supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2013, under Post-doc grant SFRH/PBD/95849/2013, by project CMUP-ERI/HCI/0051/2013 (INSIDE), and by project H2020-EU.3.7 contract 653587 (LAW-TRAIN).

References

1. Batista, F., Moniz, H., Trancoso, I., Mamede, N.J.: Bilingual Experiments on Automatic Recovery of Capitalization and Punctuation of Automatic Speech Transcripts. IEEE Transactions on Audio, Speech and Language Processing, Special Issue on New Frontiers in Rich Transcription 20(2), 474–485 (Feb 2012)
2. Campos, J., Oliveira, P., A.Paiva: Looking for Conflict: Gaze Dynamics in a Dyadic Mixed-Motive Game. Autonomous Agents and Multi-Agent Systems 30(1), 112–135 (2016)
3. Caspi, A., Harrington, H., Milne, B., Amell, J., Theodore, R., Moffitt, T.: Children's Behavioral Styles at Age 3 Are Linked to Their Adult Personality Traits at Age 26. Journal of Personality 71(4), 495–514 (2003)
4. Cloninger, S.: Conceptual issues in personality theory. In: Corr, P., Matthews, G. (eds.) The Cambridge Handbook of Personality Psychology, vol. 4, chap. 8, pp. 3–26. Cambridge Univ. Press, Cambridge, U.K. (2009)

5. Eyben, F., Scherer, K.R., Schuller, B.W., Sundberg, J., André, E., Busso, C., Devillers, L.Y., Epps, J., Laukka, P., Narayanan, S.S., Truong, K.P.: The Geneva Minimalistic Acoustic Parameter Set (GeMAPS) for Voice Research and Affective Computing. IEEE Transactions on Affective Computing 7(2), 190–202 (2016)
6. Eyben, F., Weninger, F., Gross, F., Schuller, B.: Recent Developments in openSMILE, the Munich Open-source Multimedia Feature Extractor. In: Proc. of the 21st ACM International Conference on Multimedia. pp. 835–838. New York, NY, USA (Oct 2013)
7. Kagan, J.: Three pleasing ideas. American Psychologist 51, 901–908 (1996)
8. Kelley, H., Beckman, L., Fischer, C.: Negotiating the division of a reward under incomplete information. Journal of Experimental Social Psychology 3(4), 361–398 (1967)
9. Lewis, M.: Altering fate: Why the past does not predict the future. Guilford Press, New York (1997)
10. Maćkiewicz, M., Cieciuch, J.: The Picture Based Personality Survey for Children - a new instrument to measure the Big Five in childhood. In: Proc. of the 11th Conference on Psychological Assessment. Riga, Latvia (Sep 2011)
11. Mairesse, F., Walker, M., Mehl, M., Moore, R.: Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text. Journal of Artificial Intelligence Research 30(1), 457–500 (2007)
12. Meinedo, H., Viveiros, M., Neto, J.: Evaluation of a Live Broadcast News Subtitling System for Portuguese. In: Proc. of Interspeech 2008. pp. 508–511. Brisbane, Australia (Sep 2008)
13. Mohammadi, G., Vinciarelli, A.: Automatic Personality Perception: Prediction of Trait Attribution Based on Prosodic Features. IEEE Transactions on Affective Computing 3(3), 273–284 (2012)
14. Polzehl, T., Moeller, S., Metze, F.: Automatically assessing acoustic manifestations of personality in speech. In: Spoken Language Technology Workshop. pp. 7–12. Berkeley, CA, USA (Dec 2010)
15. Polzehl, T., Moeller, S., Metze, F.: Automatically Assessing Personality from Speech. In: International Conference on Semantic Computing. pp. 134–140. Pittsburgh, PA, USA (Sep 2010)
16. Polzehl, T., Moeller, S., Metze, F.: Modeling Speaker Personality Using Voice. In: Proc. of Interspeech 2011. pp. 2369–2372. Florence, Italy (Aug 2011)
17. Rammstedt, B., John, O.: Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. Journal of Research in Personality 41, 203–212 (2007)
18. Roberts, B., Walton, K., Viechtbauer, W.: Patterns of Mean-Level Change in Personality Traits Across the Life Course: A Meta-Analysis of Longitudinal Studies. Psychological Bulletin 132(1), 1–25 (2006)
19. Schuller, B., Steidl, S., Batliner, A., Nöth, E., Vinciarelli, A., Burkhardt, F.: The INTERSPEECH 2012 Speaker Trait Challenge. In: Proc. of Interspeech 2012. Portland, OR, USA (Sep 2012)
20. Schuller, B., Steidl, S., Batliner, A., Nöth, E., Vinciarelli, A., Burkhardt, F., van Son, R., Weninger, F., Eyben, F., Bocklet, T., Mohammadi, G., Weiss, B.: A Survey on perceived speaker traits: Personality, likability, pathology, and the first challenge. Computer Speech & Language 29(1), 100–131 (Jan 2015)
21. Vinciarelli, A., Mohammadi, G.: A Survey of Personality Computing. IEEE Transaction on Affective Computing 5(3), 273–291 (2014)