Semantic Analysis of Speaker Personality Traits for Voice User Interfaces of Intelligent Systems

Veronika Krischenovich, Maksim Sokolovich, Vadim Zahariev
Belarusian State University of
Informatics and Radioelectronics
Minsk, Belarus
{krish, zahariev, m.sokolovich}@bsuir.by

Artur Rudolf
AVoice Tech LLC
Tbilisi, Georgia
arturandrudolf@gmail.com

***Abstract*—Modern speech interfaces have achieved remarkable advancements in intelligibility, naturalness of synthesized speech, and accuracy of speech recognition, enabling seamless integration into everyday life. These capabilities underpin the growing prevalence of voice-based interactions with artificial intelligence (AI) systems. However, enhancing user-machine interaction requires addressing empathy and personalization, adapting AI responses to users socio-cultural, professional, and psychological traits. This is particularly relevant as anthropomorphic robotic systems emerge, where voice may serve as the sole communication channel. This paper explores how semantic analysis of vocal characteristics can infer personal traits, leveraging open semantic technologies to improve intelligent systems adaptability and empathy.**

Speech interfaces (SI) or voice user interfaces (VUI) have evolved significantly, with text-to-speech (TTS) systems achieving high naturalness and speech-to-text (STT) systems boasting impressive recognition accuracy [1]. These advancements, driven by deep learning and large-scale datasets, have made voice a primary modality for interacting with AI, from virtual assistants like Siri and Alexa to cutting-edge models like ChatGPT with voice capabilities [2]. Currently, such systems are ubiquitous, facilitating tasks from ordinary scheduling to complex problem solving [3].

***Keywords*—semantic analysis, speech interfaces, OSTIS, personalization in AI, personality traits recognition, intelligent systems, empathy in human-machine interaction, OCEAN personality model, real-time voice processing**

I. Introduction

However, challenges remain in moving these interactions beyond the usual functionality. Users increasingly expect empathetic, context-aware responses tailored to their individuality. This demand is further amplified by the growing popularity of anthropomorphic robots, such as Atlas by Boston Dynamics, Ameca by Engineered Arts and and Optimus by Tesla, designed for human-like interaction [4]–[6]. In such systems, voice often becomes the sole input source, which requires advanced analysis to decode not only the content, but also the speaker's personal characteristics — accounting for extralinguistic features of their speech: psychological, sociocultural, professional, and other aspects of their behavior [7].

This enhances user engagement during interaction and fosters a positive overall user experience, as personalized and emotionally nuanced responses create a sense of genuine dialogue, strengthening trust and attachment to the system. When an interface not only solves a task but also considers the user's context, emotional state, and cultural background, they feel heard and understood, reducing cognitive load and increasing satisfaction. In the long term, this builds loyalty, as the interaction ceases to be transactional and evolves into meaningful, human-like communication.

II. State of the art in emphatic voice interfaces for intelligent systems

Despite this technological maturity, most speech interfaces deliver generic responses that lack personalization, a limitation evident in social robots used for therapy, as described in the study on a personalized behavior control system for social robots [8]. These robots, designed to assist children with autism, often fail to adapt to individual emotional or behavioral cues, such as when a child's heightened excitement requires a calming response, yet the system provides a standard reply, potentially disrupting engagement. A promising advancement addressing this gap is the Empathic Voice Interface (EVI), introduced by Hume AI [9]. EVI analyzes prosodic features like pitch variation and tempo to infer emotional states such as joy or stress, adjusting its tone and phrasing to enhance perceived empathy, as seen when it softens delivery for a trembling voice. Supported by findings in Frontiers in Robotics and AI (2024), which link emotionally responsive interfaces to improved trust, EVI leverages emotional speech datasets akin to those in Scientific Data (2024), marking a shift toward emotionally intelligent AI.

However, integrating such innovations into broader intelligent systems remains a challenge, as real-time personalization across socio-cultural, professional, and psychological user profiles is still nascent. The OSTIS framework, with its semantic network approach, offers a solution by decoding vocal cues-pitch, rhythm, hesitations—beyond linguistic content to infer user traits and

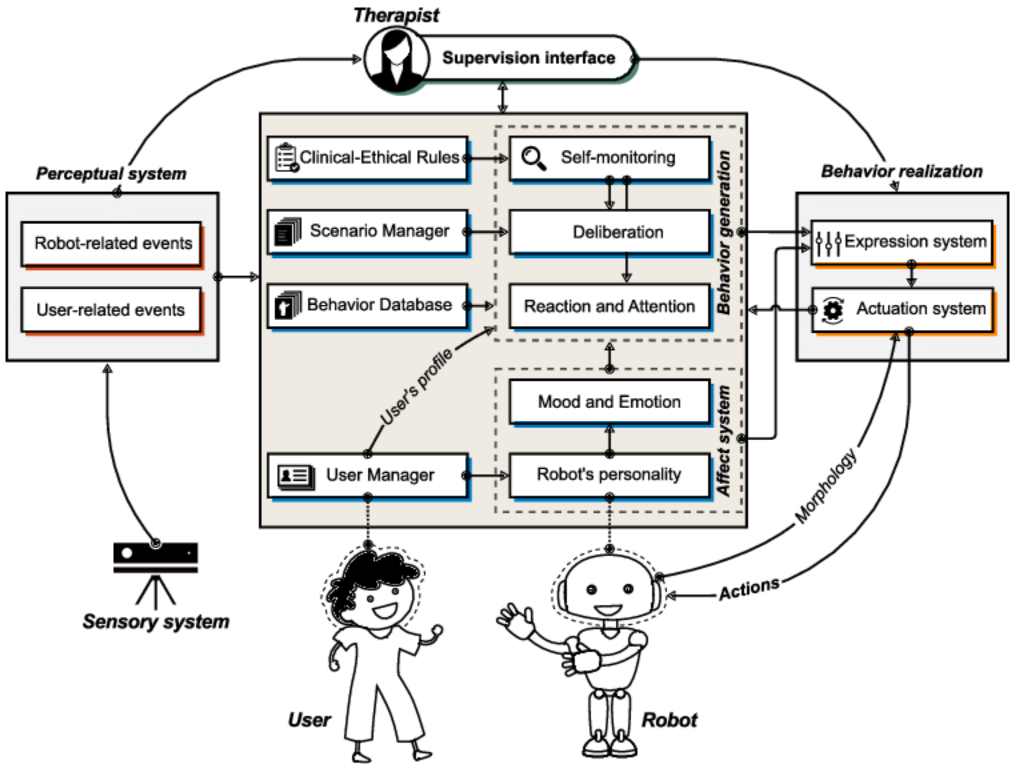

Figure 1. Example of a personalized social robotics system for therapist-assisted interaction with children with autism [8]

adapt responses, such as adopting a formal tone for professionals or a calming cadence for stressed speakers [10], [11]. This capability is critical as anthropomorphic robots and virtual assistants increasingly dominate human-AI interaction, where voice often serves as the sole input. By formalizing vocal-emotional mappings in a scalable knowledge base, OSTIS can bridge the personalization gap, moving speech interfaces beyond functionality toward deeper, trait-sensitive communication that aligns with users' unique needs and expectations.

III. Problem of personality traits formalization and modeling

Personality traits are enduring patterns of thoughts, feelings, and behaviors that distinguish individuals from one another, shaping how they interact with the world. According to Matthews, Deary, and Whiteman [12], these traits are relatively stable over time and across situations, emerging from a complex interplay of genetic, environmental, and developmental factors. They provide a framework for understanding individual differences, influencing everything from social relationships to decision-making styles. In the context of intelligent systems, recognizing these traits via voice can enable more adaptive, user-centric interactions, making their formalization a critical step for advanced speech interfaces [13].

Various models have been proposed to categorize personality traits, including the Big Five, Myers-Briggs Type Indicator (MBTI), HEXACO, Dark Triad, Eysenck's PEN model, and the 16 Personality Factors (16PF) [14]. Among these, the Big Five, HEXACO, Dark Triad, and 16PF stand out as the most scientifically verified, backed by extensive empirical research and widespread use in psychological studies. The Big Five, or Five-Factor Model (FFM), boasts over 10,000 citations in peer-reviewed literature (Costa & McCrae, 1992) [15], while HEXACO, an extension of the Big Five, gains traction for its inclusion of honesty-humility (Ashton & Lee, 2007) [16]. The Dark Triad—narcissism, Machiavellianism, and psychopathy—excels in clinical and forensic contexts with robust validation (Paulhus & Williams, 2002), and the 16PF, with decades of refinement, supports detailed profiling (Cattell, 1946). In contrast, MBTI and PEN, while popular, lack comparable empirical rigor, with fewer publications grounding their constructs in real-world data. The Big Five emerges as the most commonly used and cited system, owing to its simplicity, reliability, and cross-cultural applicability.

To evaluate the applicability of personality traits models for semantic analysis in voice user interfaces, a comparison of six prominent models is presented in Table I (where Emp. Val. – Empirical Usage, the degree to which the model is applied in professional settings for assessment, diagnosis, or therapy; Cross-Cult. – Cross-Culture, the degree to which the model's concepts and structure are consistent and applicable across different cultures). The data for this table was collected from semanticscholar.org by searching for the number of references to each model in scientific publications over the last five years in the fields of psychology and computer science. The models are assessed based on citation count, clinical usage, empirical validation, cross-cultural applicability, and simplicity, as these factors determine their suitability for intelligent systems requiring robust and adaptable frameworks.

The Big Five model stands out with over 10,000 citations [15], reflecting its widespread adoption and robust empirical support. Its simplicity and cross-cultural applicability make it ideal for general-purpose applications, including voice-based intelligent systems. HEXACO, while empirically strong, is less cited and primarily used in research due to its recent emergence [16]. The Dark Triad excels in clinical and forensic contexts [17], offering targeted insights into specific traits but lacking broad applicability. The 16PF provides detailed profiling for clinical diagnostics [18], though its complexity limits accessibility. MBTI and PEN models, despite popularity in non-clinical settings, suffer from weaker validation, reducing their reliability for advanced systems.

The Big Five model comprises five core scales, often described using the so-called OCEAN notation: Openness (creativity, curiosity, and preference for novelty), Conscientiousnes (self-discipline, organization, and dependability), Extraversion (sociability, talkativeness, and assertiveness), Agreeableness (compassion, cooperativeness, and trust) and Neuroticism (emotional instability, anxiety, and moodiness). Each of these scales represents a continuum of traits observable in behavior and, notably, vocal patterns [13]. These scales are:

- **Openness to Experience**: Reflects a person's imagination, creativity, and willingness to embrace new ideas or unconventional perspectives. High scorers are often curious and artistic, exhibiting varied pitch and expressive intonation in speech, while low scorers prefer routine and may speak in a more monotone, predictable manner.
- **Conscientiousness**: Measures organization, dependability, and goal-directed behavior. Individuals high in this trait are methodical and deliberate, often speaking with clear, steady pacing, whereas those low in conscientiousness may display erratic rhythms or frequent interruptions, signaling impulsivity.
- **Extraversion**: Captures sociability, assertiveness, and energy levels. Extroverts tend to speak loudly, quickly, and with animated prosody, reflecting their outgoing nature, while introverts use softer tones and slower tempos, indicating reserve or introspection.
- **Agreeableness**: Indicates compassion, cooperation, and likability. High agreeableness manifests in warm, gentle vocal tones and supportive phrasing, whereas low scorers may sound brusque or confrontational, with sharper intonation reflecting competitiveness or skepticism.
- **Neuroticism**: Assesses emotional instability, anxiety, and moodiness. Those high in neuroticism often exhibit tense, shaky voices or frequent hesitations under stress, while emotionally stable individuals maintain calm, even-toned speech even in challenging situations.

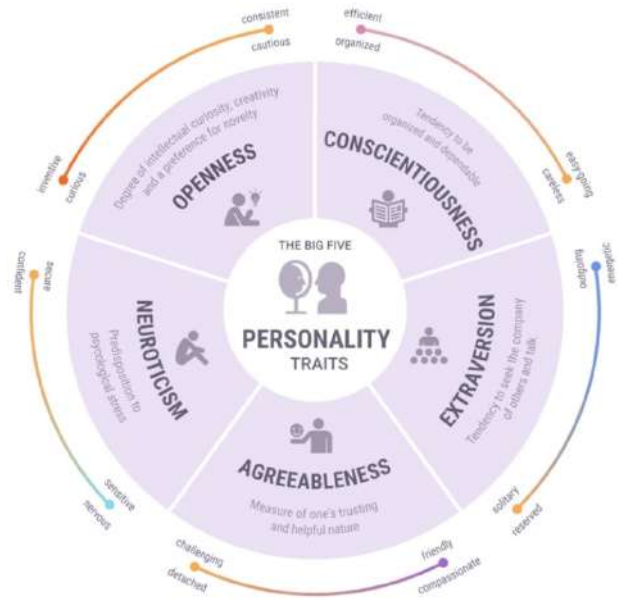

Figure 2. The Big Five personlity traits model [19]

Formalizing these scales within an ontology, such as the OSTIS framework, involves creating a structured knowledge base that maps acoustic features (e.g., pitch, tempo) to trait dimensions. However, real-world applications of models like the Big Five reveal limitations—scales often require clarification and expansion to capture a more comprehensive picture of personality traits, as their broad categories may oversimplify nuanced behaviors. Furthermore, intelligent systems benefit from combining different models (e.g., integrating HEXACO's honesty-humility with the Dark Triad's clinical insights) and flexibly adjusting to the specific needs of an application, such as mental health support versus professional training. OSTIS technology is essential here, offering a dynamic framework to refine and extend trait scales, merge diverse models, and adapt them contextually. By providing a machine-readable representation of personality, this semantic approach enables intelligent systems to systematically interpret vocal data, infer user traits, and tailor responses—e.g., slowing speech for a high-Neuroticism user to reduce anxiety. Ultimately, OSTIS enhances adaptability and empathy, laying the groundwork for next-generation speech interfaces that resonate with users' psychological profiles across varied use cases.

IV. Proposed approach

The formalization of the Big Five personality traits within the OSTIS framework offers a robust pathway to enhance speech interfaces by enabling semantic analysis of vocal characteristics. However, the challenge lies in bridging the gap between the raw physical parameters of a speech signal and high-level personality inferences. Low-level signal parameters are numerous and diverse,

Table I
Comparison of Personality Traits Models

Model	Citations	Clinical Usage	Emp. Val.	Cross-Cult.	Simplicity
Big Five (FFM)	> 10000	Moderate	Strong	High	High
HEXACO	~ 5000	Limited	Strong	Moderate	Moderate
Dark Triad	~ 3000	High	Strong	Moderate	Moderate
16PF	~ 2500	High	Strong	Moderate	Low
MBTI	~ 1500	Low	Weak	Low	High
PEN Model	~ 1000	Low	Moderate	Low	Moderate

encompassing timbre descriptors like Fourier coefficients (FC), Linear Predictive Coding (LPC), cepstral coefficients (CC), Mel-Frequency Cepstral Coefficients (MFCC), harmonic representations, and filter bank outputs, as well as prosodic features such as fundamental frequency (F0), jitter, shimmer, syllabic/phrasal tempo, tempo deviation, and rhythm. These parameters can be measured across different psychoacoustic scales — logarithmic, Mel, Bark, etc. and ussualy supplemented by first- and second-order derivatives to capture dynamic changes. The resulting feature vector for a single analysis window often spans thousands of elements, creating high-dimensional data that is computationally intensive and not always optimal for personality trait recognition.

To address this, we develop a novel approach, initially proposed by Rudolf A. A., that compresses low-level signal features into a set of 14 high-level psychometric scales, designed to reflect human perception of a speaker's voice in intuitive and semantically meaningful categories. Using an acoustic-semantic analysis approach, proposed in [20], [21], and [22], and combining these with the semantic 14 psychometric scales, we derive a robust framework for voice quality assessment. This framework bridges the gap between objective acoustic measurements and subjective human judgments, enabling interpretable and perceptually relevant voice characterization. These scales are: Tempo, Tonality, Vibration, Voice Density, Structurality, Harmonicity, Atmosphere, Deafness/Voicefulness, Voice Volume, Voice Size, Nasalisation, Prosody, Musicality, Emotionality. Each scale is measured as a percentage ranging from 0 to 100, where these edge values represent the minimum and maximum possible expression of the given perceptual trait. Unlike raw signal parameters, these categories align with how listeners qualitatively perceive vocal traits – e. g., a "dense" voice might suggest confidence, while high "emotionality" could indicate expressiveness. By reducing the dimensionality of the feature space, this approach enhances computational efficiency while preserving interpretability, making it suitable for real-time applications in intelligent systems.

These psychometric scales serve as an intermediate layer, linking physical signal parameters to the Big Five personality traits (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism) within an OSTIS ontology. Using OSTIS's SC-code, we formalize relationships between low-level features, psychometric scales, and personality traits in a hierarchical semantic network. For instance, high tempo and emotionality might correlate with Extraversion, while stable tonality and structurality could indicate Conscientiousness. The ontology explicitly represents these mappings, allowing the system to reason about vocal cues and adapt responses dynamically. Below, we illustrate a simplified SC-code representation for the trait "Extraversion" and its associated vocal characteristics:

Extraversion
$:=$ [Big Five personality trait]
$\in$ *personality trait*
$\Rightarrow$ *description**:
[A trait characterized by sociability, assertiveness, and high energy, often reflected in animated and rapid speech patterns]
$\Rightarrow$ *associated psychometric scales**:
{• *Tempo*
  $\Rightarrow$ *description**:
  [Rate of speech, measured as syllables per second, with higher values indicating faster delivery]
  $\Rightarrow$ *correlation**:
  [Positive correlation with Extraversion, as extroverts speak quickly]
• *Emotionality*
  $\Rightarrow$ *description**:
  [Degree of emotional expressiveness in voice, captured through pitch variation and intensity]
  $\Rightarrow$ *correlation**:
  [Positive correlation with Extraversion, reflecting animated prosody]
• *Voice Volume*
  $\Rightarrow$ *description**:
  [Loudness of speech, measured in decibels]
  $\Rightarrow$ *correlation**:
  [Positive correlation with Extraversion, as extroverts tend to speak louder]
}
$\Rightarrow$ *associated low-level parameters**:
{• *Fundamental Frequency (F0)*
  $\Rightarrow$ *description**:
  [Base pitch of voice, measured in

Hertz]
⇒ *measurement scale**:
[Mel scale]
⇒ *correlation**:
[Higher variability in F0 linked to Emotionality]
- *Mel-Frequency Cepstral Coefficients (MFCC)*
⇒ *description**:
[Timbre descriptors capturing spectral envelope]
⇒ *correlation**:
[Dynamic MFCC patterns contribute to Voice Volume and Tempo]
- *Jitter*
⇒ *description**:
[Cycle-to-cycle variation in F0, measured as percentage]
⇒ *correlation**:
[Moderate jitter enhances perceived Emotionality]

}
⇒ *relation to personality inference**:
{• *vocal expression of Extraversion*
⇒ *first domain**:
[Extraversion]
⇒ *second domain**:
{• *Tempo*
• *Emotionality*
• *Voice Volume*
• *Fundamental Frequency (F0)*
• *Mel-Frequency Cepstral Coefficients (MFCC)*
• *Jitter*
}
⇒ *reasoning rule**:
[High values of Tempo, Emotionality, and Voice Volume, combined with variable F0 and dynamic MFCC, indicate a high likelihood of Extraversion]
}

This SC-code snippet formalizes "Extraversion" as a personality trait, linking it to three psychometric scales (Tempo, Emotionality, Voice Volume) and three low-level parameters (F0, MFCC, Jitter). Each scale and parameter is described with its measurement properties and correlations, while the "relation to personality inference" specifies a reasoning rule for trait detection. Similar formalizations are constructed for the other Big Five traits, mapping their respective psychometric scales—e.g., Structurality and Tonality for Conscientiousness, or Vibration and Nasalisation for Neuroticism—to relevant signal features.

The OSTIS framework's strength lies in its ability to manage this complexity. By organizing knowledge hierarchically, it integrates thousands of low-level parameters into a compact set of semantic categories, which are then probabilistically linked to personality traits. This reduces computational overhead, as the system processes fewer, more meaningful features. Moreover, OSTIS supports dynamic updates, allowing the ontology to refine correlations based on new data or incorporate additional models like HEXACO for enhanced precision. For intelligent systems, this enables real-time adaptation—e.g., softening tone for a high-Neuroticism speaker or accelerating dialogue for an extrovert—fostering empathetic, personalized interactions. Ultimately, the proposed approach leverages OSTIS to transform raw vocal data into a structured, interpretable representation of personality, paving the way for advanced speech interfaces that resonate with users' psychological profiles.

V. Conclusion

This study presents a transformative approach to enhancing voice user interfaces (VUIs) through the semantic analysis of vocal characteristics to infer speaker personality traits, leveraging the Open Semantic Technology for Intelligent Systems (OSTIS) framework. By formalizing the Big Five personality traits—Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism—within a structured ontology, the proposed method establishes a robust pipeline for decoding complex vocal cues and translating them into actionable insights about a speaker's psychological profile. The OSTIS framework plays a pivotal role by organizing thousands of low-level speech signal parameters, such as fundamental frequency (F0), Mel-Frequency Cepstral Coefficients (MFCC), jitter, and prosodic features like tempo and emotionality, into a compact set of 14 high-level psychometric scales. These scales, including Voice Density, Emotionality, and Tempo, serve as an interpretable intermediate layer, reducing computational complexity while preserving the semantic richness needed for real-time personalization in intelligent systems.

The significance of this approach lies in its ability to move VUIs beyond generic, transactional interactions toward empathetic, context-aware communication that resonates with users' individual needs. For instance, by detecting high Neuroticism through vocal markers like a shaky tone or frequent hesitations, the system can adopt a slower, calming speech pattern to reduce user anxiety. Similarly, identifying Extraversion via rapid tempo and loud volume enables the system to match the user's energy with animated, engaging responses.

The flexibility of the OSTIS framework ensures scalability and extensibility, allowing the integration of additional personality models, such as HEXACO's honesty-humility dimension or the Dark Triad's clinical insights, to refine trait detection. Its hierarchical knowledge representation supports continuous updates, enabling the system to incorporate new datasets or vocal-emotional correlations as they emerge, thus maintaining relevance in a rapidly evolving field. This adaptability is critical for applications ranging from virtual assistants and customer service bots to mental health support tools and educational platforms, where personalized interaction can significantly improve outcomes. For example, in therapeutic settings, a VUI tailored to a user's emotional stability could guide conversations with greater sensitivity, while in professional training, it could adjust its tone to match the user's level of confidence or expertise.

Looking ahead, the proposed approach lays a foundation for next-generation speech interfaces that prioritize psychological nuance and cultural sensitivity. Future work could explore cross-linguistic vocal patterns to enhance global applicability, particularly for low-resource languages, or integrate multimodal cues, such as facial expressions or gestures, to complement vocal analysis. Additionally, addressing ethical considerations—such as ensuring user consent for personality inference and mitigating biases in trait detection—will be essential to maintain trust and fairness. By harnessing OSTIS's semantic capabilities, this research not only bridges the gap between raw speech data and high-level personality insights but also redefines human-machine interaction as a deeply personalized, empathetic experience.

References

[1] R. Jampala, D. S. Kola, A. N. Gummadi, M. Bhavanam, and I. Rani Pannerselvam, "The evolution of voice assistants: From text-to-speech to conversational ai," in *2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, 2024, pp. 1332–1338.

[2] J. Xu, "Evolution and challenges in speech recognition technology: From early systems to deep learning innovations," *Applied and Computational Engineering*, vol. 121, pp. 35–41, 2025. [Online]. Available: https://doi.org/10.54254/2755-2721/2025.19493

[3] A. J. James, N. D. Vangapalli, J. Siripurapu, and Y. R. Chinnamallu, "Integration of voice assistant with chatgpt and dall-e," in *2024 International Conference on Emerging Techniques in Computational Intelligence (ICETCI)*, 2024, pp. 95–101.

[4] Ameca. the future face of robotics. [Online]. Available: https://engineeredarts.com/robot/ameca/

[5] E. Ackerman and E. Guizzo, "The next generation of boston dynamics' atlas robot is quiet, robust, and tether free," *IEEE Spectrum*, vol. 24, 2016.

[6] M. Shamsuddoha, T. Nasir, and M. S. Fawaaz, "Humanoid robots like tesla optimus and the future of supply chains: Enhancing efficiency, sustainability, and workforce dynamics," *Automation*, vol. 6, no. 1, p. 9, 2025.

[7] G. A. Pires, C. Nero, T. Losik, J. Cerqueira, H. Ayanoğlu, and E. Duarte, "Exploring empathetic interactions: The impact of sound and reactions in human-robot relations among university students," in *International Conference on Human-Computer Interaction*. Springer, 2024, pp. 19–28.

[8] H.-L. Cao, G. Van de Perre, J. Kennedy, E. Senft, P. Gomez Esteban, A. De Beir, R. Simut, T. Belpaeme, D. Lefeber, and B. Vanderborght, "A personalized and platform-independent behavior control system for social robots in therapy: Development and applications," *IEEE Transactions on Cognitive and Developmental Systems*, vol. PP, pp. 1–1, 01 2018.

[9] (2025, Mar) Emphatic VUI. [Online]. Available: https://www.linkedin.com/pulse/enhancing-human-computer-interaction-introducing-hume-varun-ps-yxytc/

[10] V. Golenkov, N. Guliakina, and D. Shunkevich, *Otkrytaja tehnologija ontologicheskogo proektirovanija, proizvodstva i jekspluatacii semanticheski sovmestimyh gibridnyh intellektual'nyh komp'juternyh sistem [Open technology of ontological design, production and operation of semantically compatible hybrid intelligent computer systems]*, V. Golenkov, Ed. Minsk: Bestprint [Bestprint], 2021.

[11] (2021, Jun) OSTIS. [Online]. Available: https://github.com/ostis-dev

[12] G. Matthews, I. J. Deary, and M. C. Whiteman, *Personality traits*. Cambridge University Press, 2003.

[13] F. Metze, A. Black, and T. Polzehl, "A review of personality in voice-based man machine interaction," in *International Conference on Human-Computer Interaction*. Springer, 2011, pp. 358–367.

[14] S. Matz, Y. W. F. Chan, and M. Kosinski, "Models of personality," *Emotions and personality in personalized services: Models, evaluation and applications*, pp. 35–54, 2016.

[15] P. T. Costa Jr and R. R. McCrae, "Four ways five factors are basic," *Personality and individual differences*, vol. 13, no. 6, pp. 653–665, 1992.

[16] M. C. Ashton and K. Lee, "Empirical, theoretical, and practical advantages of the hexaco model of personality structure," *Personality and social psychology review*, vol. 11, no. 2, pp. 150–166, 2007.

[17] D. L. Paulhus and K. M. Williams, "The dark triad of personality: Narcissism, machiavellianism, and psychopathy," *Journal of research in personality*, vol. 36, no. 6, pp. 556–563, 2002.

[18] R. B. Cattell and H. E. P. Cattell, "Personality structure and the new fifth edition of the 16pf," *Educational and Psychological Measurement*, vol. 55, no. 6, pp. 926–937, 1995.

[19] D. Kelly, "Sciencepsychology forget the myers-briggs, use the big five," May 31, 2019. [Online]. Available: https://headstuff.org/topical/science/myers-briggs-big-five/

[20] V. A. Zahariev, T. Lyahor, N. Hubarevich, and E. Azarov, "Semantic analysis of voice messages based on a formalized context," in *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems]*, V. Golenkov, Ed. BSUIR, Minsk, 2019, pp. 103–112.

[21] V. Zahariev, D. Shunkevich, S. Nikiforov, and E. Aarov, "Intelligent voice assistant based on open semantic technology," in *Open semantic technologies for intelligent systems*, V. Golenkov, V. Krasnoproshin, V. Golovko, and E. Azarov, Eds. Cham. Springer, 2020, pp. 121–145.

[22] V. A. Zahariev, T. Lyahor, N. Hubarevich, and E. Azarov, "Conversational speech analysis based on the formalized representation of the mental lexicon," in *Otkrytye semanticheskie tekhnologii proektirovaniya intellektual'nykh system [Open semantic technologies for intelligent systems]*, V. Golenkov, Ed. BSUIR, Minsk, 2021, pp. 141–158.

СЕМАНТИЧЕСКИЙ АНАЛИЗ ЛИЧНОСТНЫХ ЧЕРТ ДИКТОРА ДЛЯ ГОЛОСОВЫХ ИНТЕРФЕЙСОВ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

Крищенович В., Соколович М.,
Захарьев В., Рудольф А.

В статье представлен подход к повышению адаптивности и приданию свойств эмпатичного взаимодействия голосовым интерфейсам интеллектуальных систем за счёт семантического анализа личностных черт говорящего. Используя технологию OSTIS, предложено формализовать модель «Большой пятёрки» в семантической онтологии, связывающей акустические параметры речи с психометрическими шкалами и чертами личности. Это позволяет системам в реальном времени анализировать голосовые характеристики, такие как темп, эмоциональность и тональность, и адаптировать ответы под психологический профиль пользователя. Такой подход улучшает взаимодействие, делая его более персонализированным и эмоционально отзывчивым, что особенно важно для антропоморфных роботов и виртуальных ассистентов. Применение OSTIS обеспечивает гибкость и масштабируемость, создавая основу для нового поколения речевых интерфейсов, способных учитывать индивидуальные особенности пользователей.

Received 31.03.2025