University of Bremen

Faculty of Business Studies and Economics

Essays on
Brand Personality Perception Through Voice

Dissertation for the attainment of the doctoral degree Dr. rer. pol. by the Doctoral Commission Dr. rer. pol. of the University of Bremen

submitted by

Olga Bosak

Bremen, September 25, 2024

Date of oral examination: December 17, 2024

First examiner

Prof. Dr. Maik Eisenbeiß

Second examiner

Prof. Dr. Kristina Klein

Acknowledgments

This dissertation would not have been possible without many people's support, encouragement, and dedication.

First and foremost, I would like to thank my doctoral supervisor, Prof. Dr. Maik Eisenbeiß. Maik, thank you for giving me the opportunity to pursue my PhD in your research group. My time at markstones has been highly rewarding, allowing me to contribute to engaging projects, explore new research areas, and develop valuable skills. You always took the time to discuss my progress, provide thoughtful feedback, and, at the same time, support my ideas. I am particularly thankful for the trust and freedom you granted me in our BMBF project and the completion of my dissertation, which has been instrumental in my professional and personal development. Thank you very much for your extensive support along this journey!

A heartfelt thank you also goes to my second reviewer, Prof. Dr. Kristina Klein. Kristina, thank you for your close supervision and constant support, even with spontaneous and last-minute requests. Your guiding feedback and honest opinion have always been invaluable in helping me achieve my research goals and improve the quality of my dissertation. I sincerely appreciate your guidance.

I would also like to thank Prof. Dr. Christoph Burmann. I appreciate your devoted time to my dissertation and valuable comments during our PhD seminars. Your insights have provided considerable impulse and helped to refine my work. I have always appreciated our exchanges, especially those beyond academic topics.

I thank my friends and colleagues at the markstones Institute who have accompanied me along the way. Your support, friendship, and shared conversations have continually opened new perspectives and greatly enriched my doctoral experience. I especially want to thank my colleague Maximilian for the engaging discussions about my project and beyond. Your honest opinions and helpful advice helped make my work better.

I am especially grateful to my parents, Galina and Johann, who have supported me unconditionally throughout my academic journey and all aspects of life. Your trust and constant encouragement have given me the strength always to pursue my path. Your support has allowed me to concentrate fully on my research and ambitions—спасибо за всё.

A special thank you goes to my partner, Hendrik, who has stood by me throughout my doctoral journey. Your patience, understanding, and care have helped me maintain courage, especially during challenging times. You have supported me in my academic work and have been a constant source of support and motivation. You recognized the significance of this dissertation for me and never once questioned its worth. In the final months, your contribution has been invaluable—without you, this work would not have become what it is today.

- Olga Bosak

Table of Contents

List of Figures VII
List of Tables VIII
List of Abbreviations X

Synopsis

1. Introduction 1
2. Research Issue 4
3. Research Objective and Outline 6
4. Summary of Dissertation Essays 11
   4.1. Essay 1 11
   4.2. Essay 2 17
   4.3. Essay 3 21

Synopsis References 27

Essay 1: Voices that Resonate: A Systematic Review and Future Research Agenda of Vocal Personality Perception for Brand Voice Selection

Abstract 34
1. Introduction 35
2. Theoretical Background 38
   2.1. Inferences from Voice 38
   2.2. Voice and Personality 40
   2.3. Brand Personality 41
3. Research Objective and Framework 42
   3.1. Research Questions 44
4. Method 46
   4.1. Study Selection 47
      4.1.1. Inclusion and Exclusion Criteria 47
   4.2. Data Extraction 49
5. Macro-Level Analysis 50
   5.1. Categorization and Evolution of Research on Vocal Personality Perception 50
   5.2. Summary of Extracted Data 55
   5.3. Voice Parameters Inducing Personality Perceptions 56

6. Micro-Level Analysis 59
6.1. Scope of Analysis 59
6.2. Brand Personality Perception Through Voice 60
6.2.1. Vocal Profile of Sincerity 60
6.2.2. Vocal Profile of Excitement 64
6.2.3. Vocal Profile of Competence 67
6.2.4. Vocal Profile of Sophistication 71
6.2.5. Vocal Profile of Ruggedness 73
7. Future Research Agenda 75
7.1. Personality Perception through Voice Quality Parameters 76
7.2. Role of Speaker's Gender in Vocal Personality Perception 78
7.3. Role of Culture in Vocal Perception of Personality 80
7.4. Multivariate Analysis of Voice Parameters 82
7.5. Meta-Analysis on the Vocal Perception of Personalities 83
8. Summary and Conclusion 84
8.1. Limitations 89
Essay 1 References 90
Essay 1 Appendix 105
Appendix A. Schematic Diagrams of the Organs of Speech Production 105
Appendix B. List of Excluding Search Terms 106
Appendix C. Detailed Search Settings per Database 106
Appendix D. Personality Perception of Voice Parameters of the Timing Dimension 107
Appendix E. Personality Perception of Voice Parameters of the Frequency Dimension 111
Appendix F. Personality Perception of Voice Parameters of the Amplitude Dimension 117
Appendix G. Personality Perception of Voice Parameters of the Voice Quality Dimension 119
Appendix H. Personality Perception of Linear Combination of Voice Parameters 129
Appendix I. Journal Overview 130

Essay 2: Can You Hear My Personality? A Conceptualization of a Brand Voice Based on Brand Personality

Abstract 131
1. Introduction 132
2. Background and Conceptual Development 135
2.1. Related Work and Scope of Study 136

2.2. Conceptual Framework ... 140
2.2.1. Voice Parameters for Personality Perception ... 142
3. Methodology and Data ... 149
3.1. Study Design and Sample ... 149
3.2. Brand Personality Rating ... 150
3.3. Vocal Stimuli ... 151
3.4. Acoustical Measures ... 154
4. Brand Voice Personalities ... 155
4.1. Development of Brand Voice Personality Scale ... 157
4.1.1. Step 1: Assessment of Data Quality ... 157
4.1.2. Step 2: Verification of Data Factorability ... 157
4.1.3. Step 3: Exploratory Factor Analysis ... 158
4.1.4. Step 4: Confirmatory Factor Analysis ... 160
4.1. Interpretation of Brand Voice Personality Scale ... 164
5. Vocal Perception of Brand Voice Personalities ... 166
5.1. Development of Brand Voice Personality Model ... 166
5.1.1. Multicollinearity Assessment ... 166
5.1.2. Measurement Invariance Testing ... 168
5.1.3. Structural Equation Model ... 169
5.2. Interpretation of Brand Voice Personality Model ... 174
5.2.1. Sincerity in Brand Voice ... 175
5.2.2. Sensitivity in Brand Voice ... 176
5.2.3. Excitement in Brand Voice ... 178
6. Implications ... 179
6.1. Managerial Implications ... 179
6.2. Theoretical Implications ... 182
7. Limitations and Future Recommendations ... 184
Essay 2 References ... 187
Essay 2 Appendix ... 202
Appendix A. Access to Vocal Stimulus Set of this Study ... 202
Appendix B. Summary of Pre-Test ... 203
Appendix C. Sample Characteristics: Socio-Demographic & Survey-Related Data ... 204
Appendix D. Overview of Selected & Excluded Voices from the Jena Speaker Set ... 205
Appendix E. Confirmatory Factor Analysis Aaker's (1997) Brand Personality Scale ... 207

Appendix F. Confirmatory Factor Analysis Geuens et al.'s (2009) Brand Personality Scale 208
Appendix G. Confirmatory Factor Analysis Grohmann's (2009) Brand Personality Scale .. 208
Appendix H. Correlogram of Personality Traits ... 209
Appendix I. Scree Plot of Unrotated Factors ... 209
Appendix J. Exploratory Factor Analysis (1st rotation) ... 210
Appendix K. Exploratory Factor Analysis (Final Result) ... 211
Appendix L. Assessment of Univariate & Multivariate Normality ... 212
Appendix M. Pairwise Correlation Coefficients between the Acoustical Measures ... 213
Appendix N. Initial Multicollinearity Assessment with 14 Acoustical Measures ... 214
Appendix O. Brand Voice Personality Models (BVP-Models) ... 215

Essay 3: Brand Voice Across Borders: A Comparison of German and American Brand Personality Perceptions Through Voice

Abstract ... 217
1. Introduction ... 218
2. Background ... 222
   2.1. Speech, Communication, and Culture ... 222
   2.2. Vocal Personality Perception and Culture ... 226
3. Research Framework ... 228
4. Methodology ... 232
   4.1. Participants and Procedure ... 232
   4.2. Vocal Stimulus Set ... 233
   4.3. Acoustical Measures ... 236
5. Results ... 237
   5.1. Development of the American Brand Voice Personality Scale ... 237
      5.1.1. Assessment of Data Quality and Verification of Data Factorability ... 238
      5.1.2. Exploratory Factor Analysis ... 239
      5.1.3. Confirmatory Factor Analysis ... 241
   5.2. Development of the American Brand Voice Personality Model ... 245
      5.2.1. Preparatory Calculations for Structural Equation Modeling ... 246
      5.2.2. Multigroup Structural Equation Modeling ... 247
6. Discussion ... 252
   6.1. American Vocal Brand Personality Perception ... 252
      6.1.1. Sincerity in Brand Voice ... 252

6.1.2. Sensitivity in Brand Voice ... 254
6.1.3. Excitement in Brand Voice ... 256
6.2. Cross-Cultural Evaluation of Vocal Brand Personality Perception ... 258
6.2.1. Comparison of Brand Voice Personality Scales ... 258
6.2.2. Comparison of Brand Voice Personality Models ... 261
7. Conclusion ... 267
7.1. Managerial Implications ... 269
7.2. Theoretical Implications ... 271
7.3. Directions for Future Research ... 273
Essay 3 References ... 277
Essay 3 Appendix ... 292
Appendix A. Categorization & Definition of Voice Parameters ... 292
Appendix B. Voice Recordings & Speaker's Characteristics from the Buckeye Corpus ... 293
Appendix C. Access to Vocal Stimulus Set of this Study ... 294
Appendix D. Sample Characteristics: Socio-Demographic & Survey-Related Data ... 295
Appendix E. Confirmatory Factor Analysis Aaker's (1997) Brand Personality Scale ... 296
Appendix F. Confirmatory Factor Analysis Geuens et al.'s (2009) Brand Personality Scale 297
Appendix G. Confirmatory Factor Analysis Grohmann's (2009) Brand Personality Scale ... 297
Appendix H. Scree Plot of Unrotated Factors ... 298
Appendix I. Exploratory Factor Analysis (Final Result) ... 299
Appendix J. Assessment of Univariate & Multivariate Normality ... 300
Appendix K. Pairwise Correlation Coefficients Between the Acoustical Measures ... 301
Appendix L. Consecutive Multicollinearity Assessments ... 302
Appendix M. American Brand Voice Personality Models (BVP-Models) ... 303
Appendix N. Final Confirmatory Factor Analysis Result for German Brand Voice Personality Scale ... 305
Appendix O. American & German Estimated Regression Coefficients of the Structural Equation Models for Sincerity ... 306
Appendix P. American & German Estimated Regression Coefficients of the Structural Equation Models for Sensitivity ... 307
Appendix Q. American & German Estimated Regression Coefficients of the Structural Equation Models for Excitement ... 308

List of Figures

Synopsis

Figure 1. Research Outline 7

Essay 1: Voices that Resonate: A Systematic Review and Future Research Agenda of Vocal Personality Perception for Brand Voice Selection

Figure 1. Brand Personality Framework by Aaker (1997) 42
Figure 2. Scholarly Work on Brand Personality Perception Through Voice Over Time 43
Figure 3. Research Framework of Literature Review on Vocal Personality Perception 46
Figure 4. PRISMA Flow Diagram of the Search and Selection Process 49
Figure 5. Evolution of Research on Vocal Personality Perception 54
Figure 6. Avenues for Future Research on Vocal Personality Perception 76

Essay 2: Can You Hear My Personality? A Conceptualization of a Brand Voice Based on Brand Personality

Figure 1. Conceptual Framework of Brand Voice Personality Model (BVP-Model) 142
Figure 2. Spectrogram and Oscillogram of Female and Male Utterances 144
Figure 3. Vibratory Vocal Folds Pattern of a: A) Modal; B) Breathy; C) Creaky; and D) Hoarse Voice 146
Figure 4. Age and Perceived Age of Speakers Within Vocal Stimuli 153
Figure 5. Four-Step Scale Development and Evaluation Process 156
Figure 6. Brand Voice Personality Scale (BVP-Scale) 165

Essay 3: Brand Voice Across Borders: A Comparison of German and American Brand Personality Perceptions Through Voice

Figure 1. Cultures Categorized According to: A) High/Low Context Cultures; B) Linear-active, Reactive, Multi-active Cultures 225
Figure 2. Research Framework of American English Perceptual Study 230

List of Tables

Synopsis

Table 1. Overview of Dissertation Essays 4
Table 2. Extended Overview of Dissertation Essays 10

Essay 1: Voices that Resonate: A Systematic Review and Future Research Agenda of Vocal Personality Perception for Brand Voice Selection

Table 1. Concept Matrix of Extracted Data 50
Table 2. Macro-Level Analysis of Extracted Data 55
Table 3. Definition of Voice Parameters 57
Table 4. Vocal Profiles of Brand Personality Traits Related to Sincerity 64
Table 5. Vocal Profiles of Brand Personality Traits Related to Excitement 67
Table 6. Vocal Profiles of Brand Personality Competence and Related Personality Traits 71
Table 7. Vocal Profiles of Brand Personality Sophistication and Related Personality Trait ... 73
Table 8. Male Voice Profile of Brand Personality Trait Masculinity 74

Essay 2: Can You Hear My Personality? A Conceptualization of a Brand Voice Based on Brand Personality

Table 1. Summary of Related Work on Vocal Personality Perception 137
Table 2. Voice Parameters for Personality Perception 143
Table 3. Overview of Brand Personality Traits Used in Survey 151
Table 4. Ranges and Means of Acoustic Measures of Vocal Stimuli Set “JESS” 155
Table 5. Average Variance Extracted and Squared Correlations 161
Table 6. Final Confirmatory Factor Analysis 163
Table 7. Model Fit Indices for Competing Measurement Models 164
Table 8. Conclusive Multicollinearity Assessment 168
Table 9. Measurement Invariance Testing 168
Table 10. Estimated Regression Coefficients of the Structural Equation Models for Brand Voice Personality Sincerity 170
Table 11. Estimated Regression Coefficients of the Structural Equation Models for Brand Voice Personality Sensitivity 172

Table 12. Estimated Regression Coefficients of the Structural Equation Models for Brand Voice Personality Excitement 174
Table 13. Voice Profile of Brand Voice Personality Sincerity per Gender 176
Table 14. Voice Profile of Brand Voice Personality Sensitivity per Gender 177
Table 15. Voice Profile of Brand Voice Personality Excitement per Gender 179

Essay 3: Brand Voice Across Borders: A Comparison of German and American Brand Personality Perceptions Through Voice

Table 1. Voice Parameters of Interest for American Personality Perception Study 231
Table 2. Overview of Used Brand Personality Traits in Survey 233
Table 3. Ranges and Means of Acoustical Measures of Vocal Stimuli Set “Buckeye” 236
Table 4. Assessment of Female and Male Speakers on Feminine and Masculine Traits 238
Table 5. Iterative Steps in Exploratory Factor Analysis Model Purification 241
Table 6. Overview of Conducted Reliability and Validity Tests 242
Table 7. Final Confirmatory Factor Analysis for American Brand Voice Personality Scale 244
Table 8. Model Fit Indices for Competing Measurement Models 245
Table 9. Measurement Invariance Testing 247
Table 10. Estimated Regression Coefficients of the Structural Equation Models for Brand Voice Personality Sincerity 249
Table 11. Estimated Regression Coefficients of the Structural Equation Models for Brand Voice Personality Sensitivity 250
Table 12. Estimated Regression Coefficients of the Structural Equation Models for Brand Voice Personality Excitement 251
Table 13. American English Brand Voice Profile of Sincerity 254
Table 14. American English Brand Voice Profile of Sensitivity 256
Table 15. American English Brand Voice Profile of Excitement 258
Table 16. German and American Brand Voice Personality Scales 259
Table 17. German and American Brand Voice Personality Profiles of Sincerity 263
Table 18. German and American Brand Voice Personality Profiles of Sensitivity 264
Table 19. German and American Brand Voice Personality Profiles of Excitement 266

List of Abbreviations

ABI	Acoustic Breathiness Index
AGFI	Adjusted Goodness-of-Fit Index
AI	Artificial Intelligence
AIC	Akaike Information Criterion
av.	Average
AVE	Average Variance Extracted
AVQI	Acoustic Voice Quality Index
BPS	Brand Personality Scale
BVP-Model	Brand Voice Personality Model
BVP-Scale	Brand Voice Personality Scale
C.R.	Critical Ratio
CASA	Computers as Social Actors
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
COO	Country of Origin
CPPS	Smoothed Cepstral Peak Prominence
CR	Composite Reliability
Db	Decibel
Df	Formant Dispersion
df	Degrees of Freedom
DGD	Database for Spoken German
EFA	Exploratory Factor Analysis
EFL	English as a Foreign Language
f0	Fundamental Frequency
GFI	Goodness-of-Fit Index
HNR	Harmonics-to-noise Ratio
Hz	Hertz
HZSK	Hamburg Centre for Speech Corpora
ICE	International Corpus of English
IMS	Institute for Natural Language Processing
IVR	Interactive Voice Response System
JESS	Jena Speaker Set

KMO	Kaiser-Meyer-Olkin
L2	Second Language
m	Minute
MI	Modification Index
MICASE	Michigan Corpus of Academic Spoken English
MSA	Measure of Sampling Adequacy
n	Number
NA	Not Applicable
NFI	Normed Fit Index
NLP	Natural Language Processing
PAF	Principal Axis Factoring
PCA	Principal Component Analysis
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RMS	Root Mean Square
RMSEA	Root Mean Square Error of Approximation
RQ	Research Question
S	Communality Ratio
s	Second
S.E.	Standard Errors
S-B	Satorra-Bentler
SD	Standard Deviation
SEM	Structural equation modeling
SJR	SCImago Journal Rank
SRMR	Standardized Root Mean Square Residual
st	Semitones
syl	Syllable
TLI	Tucker-Lewis Index
TTS	Text-to-Speech
USA/ the U.S.	United States of America / the United States
UX	User Experience
VIF	Variance Inflation Factor
WEIRD countries	Western, Educated, Industrialized, Rich, and Democratic Countries

Synopsis

1. Introduction

In the past, consumers have typically interacted with technological devices by physically clicking, typing, or swiping (Packard & Berger, 2024). With the recent emergence of voice artificial intelligence (AI) – AI technologies that can perform speech recognition and natural language processing (NLP) to engage in natural dialogues with users such as voice assistants (Wang et al., 2023) – there is a new shift in consumer interactions towards using voice (Packard & Berger, 2024). Nowadays, consumers talk to voice-based devices such as voice assistants to have them perform tasks such as searching for information, shopping, obtaining directions, or making reservations (Melzner, Bonezzi, and Meyvis, 2022). Interestingly, not only can the devices be controlled by the consumer's voice, but they can also "talk back" (Packard & Berger, 2024, p. 46). Therefore, voice assistants, such as Amazon's Alexa, Google Assistant, and Apple's Siri, facilitate a user experience that closely resembles human interaction due to their socially adapted behavior, personality, and capacity for independent and interactive communication (Wagner & Schramm-Klein, 2019).

Given the widespread adoption of voice technologies in households worldwide, organizations have begun to utilize and incorporate voice AI in their marketing and branding strategy. They create voice touchpoints along their customer's journey to provide personalized customer experiences that are convenient, fast, and easy to access (Kreutzer & Vousoghi, 2020; Mari et al., 2020; Steiner, 2018). For example, automobile manufacturers have integrated voice assistants into their vehicles that are capable of remembering and recalling the driver's preferences or preferred routes, such as Mercedes' "MBUX" or Volkswagen's "IDA" (Mercedes-Benz Group AG, 2022; Volkswagen Group, 2024). Additionally, virtual assistants, such as "Erica" from Bank of America, are utilized in the financial and banking sector to facilitate customer financial management through personalized voice interactions (Bank of America, n.d.). Other companies have adopted third-party voice assistants, such as Alexa or Google Assistant, rather than developing their own; Marriott has integrated Alexa into its rooms, utilizing it as a "hotel butler" (Cheng, 2018) to enhance the guest experience and Coca-Cola's voice-activated campaigns via Alexa rewarding users with coupons or free samples (Schwartz, 2020; VML, 2023).

Voice AI presents a new opportunity for marketers to shape their brand image and personality, and thus to strengthen customer relationships and build brand equity (Guha et al., 2023; H. Lee & Cho, 2020; Zoghaib, 2017). Brands enter into a dialogue with their customers, creating a new brand experience and forming new brand associations (Hörner, 2023). Progress in the development of voice-activated technology has facilitated two-way and more complex communication between consumers and brands, simultaneously enhancing naturalness and humanness (Jurafsky & Martin, 2020; V. Kumar et al., 2016). The advancement of social and natural communication is enabled not only by the growing realism of the assistants' voices (Medical Xpress, 2024; Skjegstad & Frühholz, 2024), but also by the tendency of individuals to ascribe human characteristics to computers and devices (Reeves & Nass, 1996). This anthropomorphization of voice assistants and brands offers many benefits that companies can leverage, as it positively influences product evaluations (Aggarwal & McGill, 2007), enhances product likeability (Chandler & Schwarz, 2010), and boosts brand recall, affection, and loyalty (Rauschnabel & Ahuvia, 2014).

While companies want to engage in voice marketing, determining the most appropriate voice for their brand is a major initial challenge. With voice AI, user and system interaction is usually limited to purely auditory communication, making the brand voice the primary focus of the exchange (Packard & Berger, 2024). Research in phonetics and psychology demonstrates the importance of voice in shaping speakers' perceptions of their emotional and motivational state, physiological cues, identity, and personality (Gobl & Chasaide, 2003; Krauss et al., 2002;

Schweinberger & Zäske, 2019; Zäske et al., 2020). Furthermore, following Mehrabian's 7-38-55 communication model, which explains how feelings and attitudes are perceived through communication, the verbal content of a message accounts for 7% of the total perception, the tone of voice contributes 38%, and non-verbal communication, such as body language, accounts for 55% (Mehrabian & Ferris, 1967). As visuals and, therefore, non-verbal communication is absent in voice-based AI interactions, the voice representing a brand primarily influences the perception of its personality (Klasmeyer & Sendlmeier, 1997; Zäske et al., 2020). Accordingly, those engaged in voice marketing activities must identify a voice that fits their brand so that customers can perceive the intended brand personality in auditory communication. Despite the importance of selecting an appropriate brand voice, marketers typically rely on their intuition and experience when choosing a brand voice (Dahl, 2010; Wiener & Chartrand, 2014). While there is a substantial corpus of empirical evidence on the perception of human personalities through voice, there is no systematic approach, methodology, or model to identify the appropriate voice for conveying the personality of a brand.

This dissertation addresses the need for research on the vocal perception of brand personalities, given the prevalence of voice-based technology and the resulting purely auditory communication between brands and their customers. Since the brand voice conveys its personality through voice-based communication, marketers require guidance in identifying a voice that aligns with their brand personality. Three consecutive essays examine how brand personalities can be perceived through voice alone, employing qualitative and quantitative methodologies. **Table 1** presents a short overview of the three dissertation essays. The following sections provide a more detailed overview of the underlying research issue and objective and clarify how each dissertation essay contributes to finding the appropriate voice for brands.

Table 1. Overview of Dissertation Essays

Essay	Title	Author
1	Voices that Resonate: A Systematic Review and Future Research Agenda of Vocal Personality Perception for Brand Voice Selection	Olga Bosak
2	Can You Hear My Personality? A Conceptualization of a Brand Voice Based on Brand Personality	Olga Bosak
3	Brand Voice Across Borders: A Comparison of German and American Brand Personality Perceptions Through Voice	Olga Bosak

2. Research Issue

Research on personality perception through voice has existed for nearly a century, producing numerous concepts and theories (Allport & Cantril, 1934; Moore, 1939; Pear, 1931). The emergence of innovative digital technologies and software designed to enhance speech processing analysis, along with advancements in phonetics and psychology, led research on voice perception to receive considerable attention between the 1960s and 1980s (e.g., Addington, 1968; Apple et al., 1979; Brown et al., 1985; Scherer, 1972). This extensive research on vocal personality perception showed that especially the paraverbal content of speech significantly influences the perception of personality traits, such as competence, dominance, and extraversion (Kreiman & Sidtis, 2011; Mehrabian & Ferris, 1967).

Paraverbal content refers to the acoustic parameters of the voice surrounding the semantic content of a message, such as speech rate, pitch, or intensity (Ketrow, 1990). All humans share a comparable structure of the vocal tract system. Yet, differences in the length of vocal folds and slight variations in the acoustic parameters around a mean contribute to the unique characteristics of each voice (Belin et al., 2011; Rodero, 2013). Despite such voice uniqueness, psychologists have discovered correlations between several acoustic parameters and specific personality traits. Voice stereotypes are said to exist, meaning that single acoustic parameters or combinations of them induce the same personality to be perceived by listeners of various voices (Addington, 1968; Kramer, 1964).

The research into the vocal perception of personalities conducted over several decades has increased its utilization in marketing and consumer research. At a time when telemarketing was a common practice, it was found that interviewers who exhibited a high pitch, great pitch variability, loudness, and a fast speech rate were perceived as competent and credible and, therefore, had a lower rejection rate (Ketrow, 1990; Oksenberg et al., 1986; Sharf & Lehman, 1984). Furthermore, research has demonstrated a significant correlation between a fast speaking rate and enhanced sales performance in direct sales settings (Peterson et al., 1995). In radio commercials, a voice that fits the advertised brand improves product recall and increases the willingness to buy the product (Dahl, 2010; North et al., 2004). Nevertheless, despite the significance of voice in marketing, the existing research on vocal personality perception has exclusively focused on investigating human personalities. Research on the vocal perception of brand personalities is missing.

Brand personalities are conceptualized as brands possessing humanlike traits and were found to function the same way human personalities do (Wee, 2004). Yet, studies on the compatibility between human and brand personalities have also demonstrated that not all personality traits that describe a human can be transferred to brands (A. Kumar, 2018). One possible reason for this discrepancy is that personality traits applied to humans may have different meanings when applied to brands (Caprara et al., 2001). Also, some human personality traits do not apply to brands; for instance, human traits like “neurotic fatigue” (Azoulay & Kapferer, 2003, p. 149). Further, brands are more likely to be associated with excessive characteristics such as “sophisticated” or “rugged,” which humans desire yet do not possess (Aaker, 1997).

Thus, although the concepts of human and brand personalities appear to be similar, it is necessary to differentiate between personality traits that are applicable to brands and those that are not. Consequently, further research is required on the vocal perception of relevant personalities that are applicable to brands.

3. Research Objective and Outline

Reacting to the aforementioned research gap, the overall objective of this dissertation is to create insights into the perception of brand personalities through voice. Being the first to study this topic, the findings significantly contribute to the field of research and advance the knowledge about the vocal perception of brand personalities. Further, the dissertation's results guide marketers and voice specialists in determining, selecting, or creating the appropriate voice to fit the desired personality of a brand. Regarding the increasing significance of voice AI integration in marketing and branding strategy, the insights and results of this dissertation help to overcome the managerial challenges that come with it.

The present dissertation is specifically concerned with the perception of brand personalities conveyed through human voices, as opposed to those of artificial sources. Empirical evidence has shown that human voices are preferred in human-computer interactions over their artificial counterparts (Atkinson et al., 2005; Chiou et al., 2020; Seaborn et al., 2022; Xu, 2019). This preference is based on the assumption that natural voices are more effective, elicit greater attention, and facilitate better recall with less concentration (Atkinson et al., 2005; Rodero, 2017).

This dissertation investigates the vocal perception of brand personalities from different perspectives. It begins with a qualitative study in Essay 1, which represents the current state of research in the vocal perception of human personalities and derives a future research agenda. Essays 2 and 3 deal with the vocal perception of brand personalities in an empirical manner. **Figure 1** shows the research outline and illustrates how the three essays in this dissertation are interconnected and deal with different but complementary aspects of the topic of vocal brand personality perception.

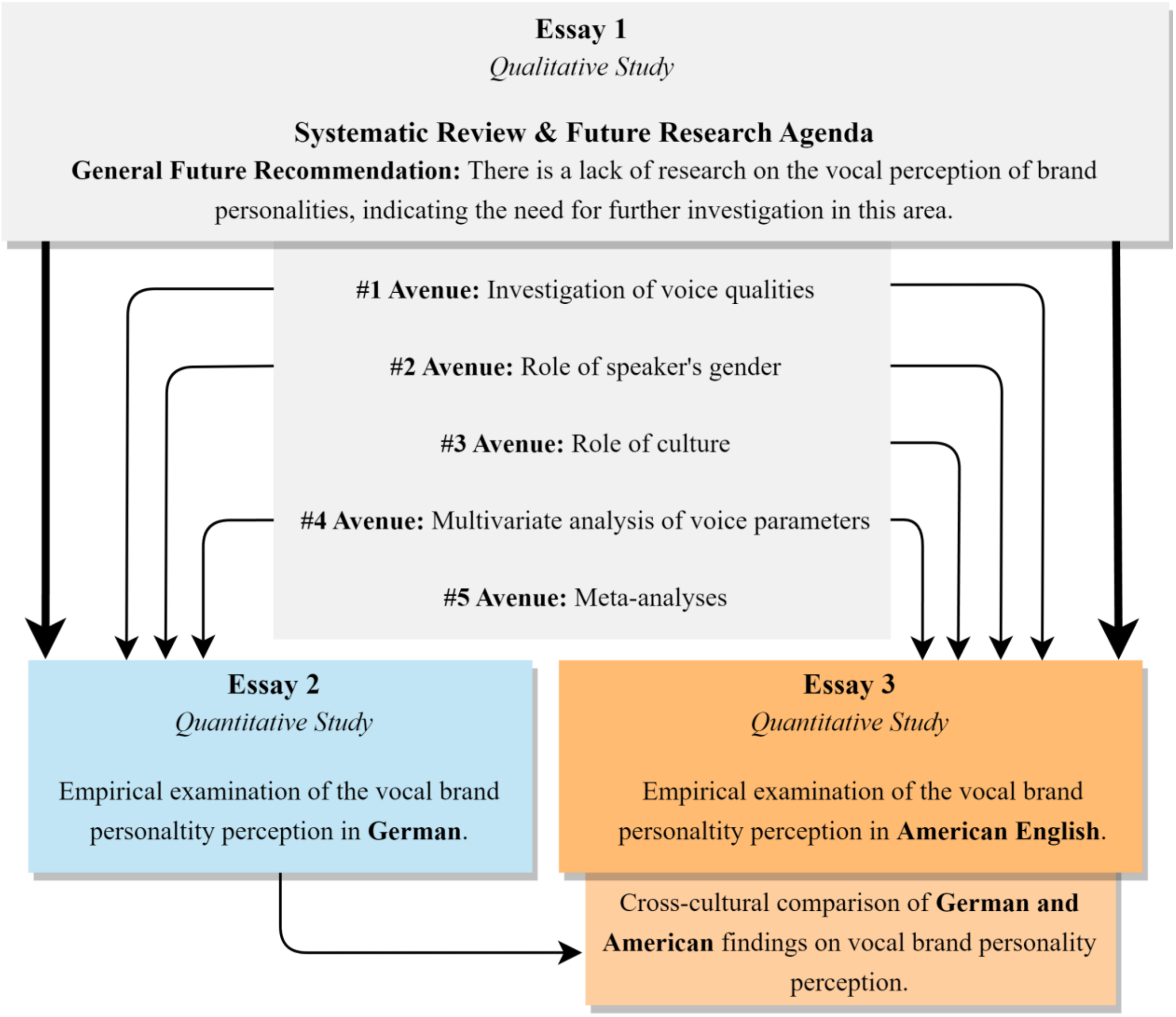


Figure 1. Research Outline

In Essay 1, titled "Voices that Resonate: A Systematic Review and Future Research Agenda of Vocal Personality Perception for Brand Voice Selection", the existing research on vocal human personality perception is identified, summarized, and classified through a systematic literature review. In light of the empirical findings on human personalities, conclusions are drawn regarding the vocal perception of the brand personality dimensions of sincerity, excitement, competence, and sophistication. These brand personalities are derived from the first developed brand personality scale (BPS) and are of particular interest as they are widely known and applied by academia and practitioners to strategically shape and communicate the brand identity (Aaker, 1997). The review of Essay 1 represents the first comprehensive systematic review of human personality perception through voice parameters, covering nearly 85 years of published research. Further, it guides marketers in selecting an appropriate brand voice through the derivation of vocal profiles for sincere, exciting, competent, and sophisticated brands.

The comprehensive review of the literature on vocal personality perception in Essay 1 provides a foundation for the following two essays, demonstrating that empirical research on the vocal perception of brand personalities is generally missing. The studies in Essays 2 and 3 respond to this gap in research by empirically investigating how brand personalities are perceived through voice. Moreover, the review's findings and identified research gaps are utilized to develop a future research agenda for vocal brand and human personality perception. Essays 2 and 3 respond to four of the five recommended future research avenues: both essays utilize multivariate analysis of voice parameters, investigate the voices of females and males, and incorporate the analysis of objectively measurable voice qualities. Additionally, Essay 2 focuses on brand personality perception in German, while Essay 3 examines brand personality perception in American English with a cross-cultural comparison of the results.

Essay 2, titled "Can You Hear My Personality? A Conceptualization of a Brand Voice Based on Brand Personality", investigates how brand personalities are perceived through voice by developing a brand voice personality scale (BVP-Scale) and a brand voice personality model (BVP-Model). The BVP-Scale measures the brand personalities identified to be perceived by voice alone, namely sincerity, sensitivity, and excitement, which are defined as the brand voice personality dimensions. The BVP-Model indicates the precise linear combination of objectively measurable voice parameters resulting in the perception of the identified brand voice personality dimensions per voice gender. The exploratory study of Essay 2 represents the first empirical study to examine the perception of brand personalities through voice, given that previous perceptual studies have primarily focused on human personalities. The focus of Essay 2 is on a specific language area, using German voices and brand personality traits. However, since individuals learn to utilize their voices in culturally determined ways (Krauss et al., 2002; Waaramaa et al., 2021), differences in brand personality perceptions are expected to emerge across countries. Thus, it is recommended that the study be replicated in English-speaking countries and countries with significantly different communication models.

Therefore, Essay 3, titled “Brand Voice Across Borders: A Comparison of German and American Brand Personality Perceptions Through Voice”, replicates the German study of Essay 2 in an American English-speaking context to ascertain the cross-cultural validity of the findings. Following the German approach, the study first determines which brand personalities can be perceived through voice in an American English-speaking context by developing an American BVP-Scale. Further, it develops an American BVP-Model that determines how female and male voices induce the perception of brand voice personality dimensions. Building on the findings, Essay 3 focuses on determining the cultural differences in vocal brand personality perceptions between the United States and Germany. Therefore, a comparison of the German and American BVP-Scales and BVP-Models is made. In doing so, the replication study of Essay 3 represents the first cross-cultural investigation of the vocal perception of brand personalities. Please refer to **Table 2** for an extended overview of the dissertation essays.

Table 2. Extended Overview of Dissertation Essays

Essay	Title	Publication Status	Research Objectives	Data	Key Findings
1	Voices that Resonate: A Systematic Review and Future Research Agenda of Vocal Personality Perception for Brand Voice Selection	In preparation for submission to the Journal of Product & Brand Management.	1. Determining which voice parameters induce human personality perceptions to derive inferences on the vocal perception of **brand personalities**. 2. Providing a **managerial guide** on selecting a voice that fits a desired brand personality. 3. Developing a **future research agenda** for vocal brand and human personality perception.	• Systematic literature review of 52 articles on vocal personality perception from 1939 – 2023 from interdisciplinary journals.	• **First comprehensive summary** of 444 single results on the vocal perception of 59 human personalities through 22 voice parameters. • Derivation of **male and female vocal profiles** for the brand personality dimensions of sincerity, excitement, sophistication, and competence to guide marketers in brand voice selection. • **Future agenda with five research avenues**: voice qualities, role of speaker's gender, role of culture, multivariate analysis of voice parameters, and meta-analysis.
2	Can You Hear My Personality? A Conceptualization of a Brand Voice Based on Brand Personality	A preliminary version was presented at the EMAC Conference 2023 (see Bosak et al., 2023).	1. Developing a **model** that determines how brand personalities are perceived through voice. 2. Identifying which **brand personalities can be perceived through voice**. 3. Determining which linear combinations of objectively measurable **voice parameters induce brand personality perceptions**.	• Pre-test online survey: N = 33 • Main study online survey: N = 2,000 • Vocal stimulus set: 96 voices from Jena Speaker Set (JESS) • German sample and vocal stimuli	• **First empirical study** to examine the perception of brand personalities through voice. • Development of a **brand voice personality scale (BVP-Scale)** consisting of three dimensions: sincerity, sensitivity, and excitement. • Development of a **brand voice personality model (BVP-Model)** that indicates the precise combination of voice parameters resulting in the perception of a specific brand voice personality dimension per gender.
3	Brand Voice Across Borders: A Comparison of German and American Brand Personality Perceptions Through Voice	Essays 2 and 3 will be combined into one article and submitted to a leading marketing journal. Preparations for the submission are currently being undertaken.	1. Replicating the German study of Essay 2 in an American English context. Thus: a. Developing an **American BVP-Scale**; b. Developing an **American BVP-Model**. 2. **Cross-cultural comparison** of German and American BVP-Scales and BVP-Models.	• Online survey: N = 979 • Vocal stimulus set: 40 voices from Buckeye corpus • U.S. sample and vocal stimuli	• Development of an **American BVP-Scale** consisting of three dimensions: sincerity, sensitivity, and excitement. • Development of an **American BVP-Model** that determines how female and male voices induce the perception of brand voice personality dimensions. • **First cross-cultural study** investigating the vocal perception of brand personalities in Germany and the United States.

4. Summary of Dissertation Essays

This section comprises a comprehensive summary of each of the three dissertation essays. Each summary outlines the respective essay's motivation, objectives, methodology, key findings, and contribution to the literature on vocal personality perception and voice marketing.

4.1. Essay 1

"Voices that Resonate: A Systematic Review and Future Research Agenda of Vocal Personality Perception for Brand Voice Selection"

Author: Olga Bosak

The interactions between consumers and voice-based devices illustrate an emerging paradigm shift in consumer-brand communication. Whereas traditionally, brand interactions have mainly been one-way, voice technology allows for bidirectional, sophisticated, and increasingly human-like conversations (Jurafsky & Martin, 2020; V. Kumar et al., 2016; Packard & Berger, 2024). The advantages of voice-based communication are manifold; however, the advent of voice technology also presents a challenge for marketers in maintaining brand conversations with customers, given that voice devices and platforms typically lack visual content (Packard & Berger, 2024). As a result, brands rely solely on auditory communication, making the voice the primary focus of the exchange. It is, therefore, crucial to select an appropriate and distinctive voice for an organization and its brand, one that conveys a desired attitude and personality to enhance the customer experience and differentiate the brand from competitors (McAleer & Belin, 2019).

Despite the substantial corpus of empirical evidence on vocal personality perception and its marketing implications, there is a lack of a comprehensive summary, overview, or guide specifying which voice characteristics lead to personality perceptions. In addition, perceptual studies have mainly focused on the vocal perception of human personalities, not brand personalities. Brand personalities are conceptualized as brands possessing humanlike traits and were found to function the same way human personalities do (Wee, 2004). However, research on the compatibility between human and brand personalities has also revealed that not all personality traits that describe a human can be transferred to brands (A. Kumar, 2018). Therefore, research on the vocal perception of relevant personalities applicable to brands is needed. Consequently, a managerial guide on selecting an appropriate voice to reflect a desired brand personality is currently missing, so marketers predominantly rely on their gut feelings or experiences to choose a brand voice (Dahl, 2010; Wiener & Chartrand, 2014).

Essay 1 addresses this issue by identifying, summarizing, and classifying existing research on vocal human personality perception through a systematic literature review. The study aims to determine which voice parameters induce human personality perceptions to derive inferences on the vocal perception of brand personalities. In doing so, the study follows the first developed and widely-known dimensions of Aaker's (1997) BPS, sincerity, excitement, competence, sophistication, and ruggedness, and derives vocal profiles for these brand personality dimensions based on the review results. In addition to this primary research objective, the review is guided by further research questions concerning the differences in stimulus material (speaker's gender, language, speech data type) and methodology (study design, assessment method) of perceptual studies. Essay 1 examines how the literature on vocal personality perception addresses these aspects to formulate managerial recommendations for selecting the appropriate voice to fit a brand personality. Further, the findings and identified research gaps are used to derive a future research agenda for vocal brand and human personality perception.

Due to the interdisciplinary character of the research on the vocal perception of personality, literature was searched in scientific databases, such as Scopus, Web of Science, ProQuest, and PsycNet, which provided access to various online citations and literature. The study selection process yielded 52 articles from 1939 – 2023 for review. The selected articles for the review needed to meet the following criteria: 1) they should include quantitative empirical findings in which significant correlations between voice parameters and human personality traits were found; 2) they should be published in a peer-reviewed journal; and 3) they should be written in English. Following the research framework, the data extracted from the reviewed perceptual studies was classified in a concept matrix according to the findings on the (combinations of) voice parameters that induce human personality perceptions, the stimulus material used, and the methodological aspects employed. Based on this categorization, the empirical findings of the studies were synthesized for the analyses and interpretations on a macro and micro level.

The identified articles and the extracted data on vocal personality perception are presented descriptively in the macro-level analysis. Literature on personality perception through voice is clustered into four periods highlighting the advancements in phonetic technology and the evolving interests within vocal personality perception research: 1st period of foundation: 1939 – ’70; 2nd period of cue synthesis studies: 1971 – ’80; 3rd period of cross-cultural studies: 1981 – ’99; and 4th period of contemporary research: since 2000. Further, the reviewed studies provided 444 single results on the relationship between 22 voice parameters and 59 personalities. These results are presented in a summary format and classified according to the concept matrix, providing the first comprehensive overview of the vocal perception of personalities. Finally, at the macro level of analysis, the identified 22 voice parameters, such as speaking rate, pitch, loudness, breathiness, tension, and creakiness, are defined and explained, as their understanding is essential for interpreting the findings.

In the subsequent micro-level analysis, the findings on which (combinations of) voice parameters induce personality perceptions are analyzed and interpreted. In doing so, the results on the vocal perception of the identified 59 human personalities are screened for applicable personalities to brands, following the brand personality dimensions of Aaker (1997): sincerity, excitement, competence, sophistication, and ruggedness. Therefore, drawing on the findings of the literature on vocal personality perception, male and female vocal profiles for sincere, exciting, competent, and sophisticated brands are presented, which shall guide marketers in selecting the appropriate voice for their brand personality:

- Brand personality **sincerity** is perceived when *male voices* with a high pitch variability or low loudness are used. Voice qualities like sharpness, thinness, tension, nasality, or flatness mitigate the perception of sincerity-related personality traits. Brands with *female voices* should use a voice with high pitch variability, breathiness, tension, or less resonance to communicate sincerity. Furthermore, a medium speaking rate is preferable, and vocal qualities such as nasality or roughness should be avoided.
- Brand personality **excitement** is perceived when *male voices* with a high pitch, high loudness, or a dynamic intonation are used. Additionally, the voice should be resonant or breathy but less rough or tense. Exciting brands with *female voices* should be loud or have a dynamic intonation. Additionally, an increased speaking rate, breathiness, tension, or resonance improves the perception of excitement. Voices with roughness, nasality, or creakiness should be avoided.
- The brand personality **competence** is perceived when a *male voice* with high loudness, high pitch variability, low pitch, or a fast speaking rate is selected. The linear combination of high pitch variability and a fast speaking rate increases competence perception, whereas a low pitch variability and a slow speaking rate leads to the opposite. *Female voices* should also be high in loudness, high in pitch variability, or should have a fast speaking rate. A creaky voice should be avoided, as creakiness diminishes the perception of competence.
- The brand personality **sophistication** is perceived through resonant, rough, or less nasal *male voices*. For *female voices*, results indicated that various voice qualities like

breathiness, nasality, roughness, tension, and thinness decrease the perception of sophistication.

In analyzing the findings to derive the individual vocal profiles, it becomes evident that no direct results within the reviewed articles exist for the brand personality dimensions of sincerity and excitement. Therefore, the vocal profiles for these brand personality dimensions are presented based on related personality dimensions and traits, such as agreeableness, benevolence, credibility, politeness, sensitivity, and social attractiveness, as sincerity-related personalities. Moreover, for the vocal profile of a rugged brand, “masculinity” is the only ruggedness-related personality trait found within the extracted data, which is why managerial recommendations for selecting an appropriate voice for rugged brands are impossible to make.

The identified research gaps within the literature on vocal perception of human personality are used to propose a future research agenda. The most significant future recommendation is that more research is generally needed into the vocal perception of brand personalities, especially for the brand personality dimensions of sincerity, excitement, and ruggedness. Besides this general lack of research on the perception of brand personality through voice, five avenues for future brand and human personality research are derived in Essay 1.

#1 Avenue: Future research on the vocal perception of personality should investigate the role of voice qualities since these parameters contribute significantly to defining the unique auditory color of a speaker’s voice, which in turn influences the perception of personality (Pearsell & Pape, 2023).

#2 Avenue: Studies on vocal personality perception should employ a more balanced approach by including voices of both genders, rather than focusing exclusively on either male or female voices. In particular, research on the perception of female voices is currently underrepresented, thus representing a significant gap in the existing literature that must be addressed.

#3 Avenue: Research on vocal personality perception in various languages and cultures is needed since there is a surplus of American English studies. Given the impact of cultural norms on the way individuals use their voice, which in turn affects how their personalities are perceived through the voice (Krauss et al., 2002; Waaramaa et al., 2021), further studies conducted in different languages, in addition to cross-cultural studies, are required.

#4 Avenue: In perceptual studies, voice parameters should be examined in combination, i.e., through multivariate analysis, rather than in isolation, i.e., through univariate analysis. This is because voice parameters tend to correlate in natural speech (Apple et al., 1979).

#5 Avenue: Meta-analyses on the vocal perception of personalities should be performed to provide in-depth, more accurate, and quantitative insights into how voice parameters affect personality perception.

In summary, Essay 1 contributes to the literature on vocal personality perception and voice marketing in three ways. First, it represents the first comprehensive systematic review of human personality perception through voice parameters, covering nearly 85 years of published research. Second, it shows how findings on vocal perception of human personality can be applied to brand personalities by deriving vocal profiles for sincere, exciting, competent, and sophisticated brands to guide managers in selecting brand voices. Third, next to the general need for research on the vocal perception of brand personalities, it outlines five research avenues based on current advances, such as developments in speech processing and analysis technology and the changing role of genders in society.

4.2. Essay 2

“Can You Hear My Personality? A Conceptualization of a Brand Voice Based on Brand Personality”

Author: Olga Bosak

The emergence of voice AI presents a valuable opportunity for marketers to shape their brand image and personality, thus strengthening customer relationships and building brand equity (Guha et al., 2023; H. Lee & Cho, 2020; Zoghaib, 2017). In customer-brand interactions via voice-based technology and devices, the spokesperson’s voice serves to convey the personality of the brand (Klasmeyer & Sendlmeier, 1997; Zäske et al., 2020). But how does a voice have to sound for a desired brand personality to be perceived by users? Marketers attempting to answer this question typically rely on their intuition and experience when selecting a spokesperson’s voice to represent their brand (Dahl, 2010; Wiener & Chartrand, 2014). While considerable empirical evidence exists on the perception of human personalities conveyed through voice, no methodical approach has been established for identifying the optimal voice for representing a brand’s personality.

Building on the existing body of perceptual research, Essay 2 addresses this gap by suggesting a brand voice personality model (BVP-Model) that determines how brand personalities are perceived through voice. Therefore, the initial research question concerns identifying which brand personalities can be perceived through voice alone. In doing so, a brand voice personality scale (BVP-Scale) is developed to measure the brand personalities that are perceived by voice. In light of these findings, the second research question concerns determining which combinations of voice parameters in female and male voices induce the perception of the identified brand voice personalities to form the final model.

Essay 2 begins with a review of existing literature on vocal personality perception, defining the scope of the study. Accordingly, the study employs a correlational design and utilizes external judgments of brand personalities from naïve listeners. To contribute to the existing body of research on languages other than American English, Essay 2 focuses on German voices with the intention of replicating the study in different countries in the future. Moreover, the study examines only the interaction effects of objectively measurable voice parameters and their relationship to personality judgments to ensure the results are generalizable and comparable. Additionally, as the perception of personality through voice depends on the speaker's gender, the study considers female and male voices separately to test for gender-specific effects on brand personality perception.

An online survey was conducted to pursue the research objective to develop the BVP-Model, which determines how brand personalities are perceived through voice. This survey invited listeners to evaluate two voices on the extent to which they associated the voice with the provided brand personality traits. Before this main study, a pre-test indicated that a maximum of two voices could be rated before participants demonstrated decreased attention, concentration, and willingness for voice evaluations. For the brand personality perception rating, the German translations of brand personality traits taken from the three generalizable BPS of Aaker (1997), Geuens et al. (2009), and Grohmann (2009) were used. Further, voices from the Jena Speaker Set (JESS) were selected as suitable vocal stimuli (Zäske et al., 2020). The survey yielded 3,945 individual voice ratings from 2,000 participants for the final data set (1,945 participants with two voice ratings and 55 participants with single voice ratings; $M_{age} = 51$; 53% female).

Following the initial research question, the study first identifies which brand personalities can be perceived through voice alone, i.e., brand voice personalities. To this end, a scale is developed based on a series of factor analyses and statistical procedures to explore the best model fit and establish important validities. In conclusion, the BVP-Scale is a reliable and valid measurement tool, as evidenced by its satisfactory performance on a range of psychometric properties, including indicator and construct reliability, internal consistency, discriminant and convergent validity. The developed BVP-Scale represents a reflective first-order model with three brand voice personality dimensions and eleven brand voice personality traits. The dimensions are labeled and defined as follows:

1) **Sincerity**: This dimension describes a brand voice personality perceived as honest, reliable, and sincere.
2) **Sensitivity**: This dimension describes a brand voice personality perceived as smooth, fragile, sentimental, and sensitive.
3) **Excitement**: This dimension describes a brand voice personality that is perceived as spirited, adventurous, daring, and exciting.

Following the second research question, the study determines which linear combinations of objectively measurable voice parameters induce brand voice personality perceptions, forming the final model. Accordingly, the BVP-Model is constructed based on a conducted structural equation model (SEM) with the speaker's gender-specific effects, in which the perceptions of the three brand voice personality dimensions, sincerity, sensitivity, and excitement, are related to eleven acoustical measures.[1] The acoustical measures of interest in Essay 2 derived from previous findings in psychology and phonetics regarding correlations between personality traits. In accordance with the BVP-Model, voice profiles are derived for each brand voice personality dimension providing an overview of the combination of acoustic measures that is most favorable for the respective personality perception in female and male voices:

1) **Sincerity** in brands is perceived when *females* speak fluently and loudly in brightness and show creakiness and breathiness. Likewise, a creaky *male voice* with brightness and fluent speech relates to a sincere brand personality.

[1] speaking rate, average silent pause duration, articulation rate, intensity variability, mean and range of the fundamental frequency (f0), h-h2, spectral slope and tilt, smoothed cepstral peak prominence (CPPS), and shimmer

2) **Sensitivity** in brands is perceived when *females* speak fluently, slowly, and monotonously. Further, the voice should be low-pitched and breathy but not rough. A sensitive brand represented by a *male voice* should speak brightly with breathiness and creakiness. Roughness in male voices decreases sensitivity perceptions and should be avoided.
3) **Excitement** in brands is perceived when *female* and *male voices* are breathy and creaky but not rough. Moreover, a less loud voice in terms of brightness and a less fluent speaking style with longer average silent pauses induce excitement perceptions.

The BVP-Scale and BVP-Model, combined with the derived brand voice profiles for sincerity, sensitivity, and excitement, shed light on how brand personality can be translated into a voice. Consequently, the findings of Essay 2 present marketers, sound designers, and voice engineers with a systematic methodology for selecting or designing a brand voice that fits a desired brand personality. Moreover, the managerial implications of Essay 2 highlight the significance and benefits of aligning a brand's communication with its appropriate brand personality, as conveyed through the voice of the brand. Furthermore, it outlines the general opportunities offered by voice AI and illustrates its exemplary application areas.

In conclusion, Essay 2 makes three contributions to vocal personality perception and voice marketing research. Primarily, it is the first empirical study to examine the perception of brand personalities through voice compared to previous research focusing on vocal human personality perception. Secondly, Essay 2 examines the influence of linear combinations of voice parameters on brand personality perceptions, thereby reacting to the call for multivariate analyses of voice parameters in perceptual studies (Feinberg et al., 2005; McAleer et al., 2014; Pisanski & Rendall, 2011; Wu et al., 2021). Finally, the study's exploratory approach considers only objectively measurable voice parameters and their acoustical measures, ensuring the comparability of results and expanding the knowledge on voice perceptions.

4.3. Essay 3

“Brand Voice Across Borders: A Comparison of German and American Brand Personality Perceptions Through Voice”

Author: Olga Bosak

When it comes to conversational agents and voice assistance, international acting companies must comprehend the noteworthy influence of culture as an acoustic cue and consider adapting their brand’s voice to align with the cultural context in which they wish to interact with customers. The way a brand is represented through its voice influences the perception of the brand personality and, therefore, the brand identification (Nam et al., 2011). This perception depends not only on the voice and its acoustic characteristics but also on the cultural context of the customer (Krauss et al., 2002).

Cross-cultural perceptual studies have revealed the influence of sociocultural factors on the perception of personalities through voice (H. O. Lee & Boster, 1992; Peng et al., 1993; van Bezooijen, 1995). Vocal stereotypes can vary between cultures due to differences in how personality traits are learned and perceived from speech (Kreiman & Sidtis, 2011; Waaramaa et al., 2021). For instance, Japanese cultural norms dictate that women speak more politely than men. A higher pitch in the voice conveys this politeness. Although women naturally have medium-pitched voices, cultural norms have conditioned Japanese women to speak in a higher pitch (Krauss et al., 2002). In a communication system that relies solely on voice, it is therefore crucial for organizations to understand the nuances of a culture, mainly when the objective is to enhance the user experience (UX) and build meaningful customer relationships internationally (Pang, 2021).

Despite efforts to provide culturally sensitive voice assistants in various languages, there is a limited body of research and adaptation in personality perception across languages. Although Alexa can imitate human emotions through excited or disappointed voice tones (available in British English and Japanese), it is not yet possible for any voice assistant to reflect a specific personality by adjusting its voice (Wenden, 2020). Building on previous research examining the vocal perception of personalities in different language contexts, this study investigates the role of culture in the perception of brand personality through the voice. To this end, first, perceptual research on the vocal perception of brand personality is conducted in the United States. Subsequently, the findings are compared with the results of the preceding German research presented in Essay 2.

Given that the United States represents one of the largest markets for smart speakers, it is imperative to research the perception of brand personality in American English. However, such research currently needs to be completed. By conducting a perceptual study in American English, Essay 3 follows the recommendations of the two preceding essays on cross-cultural research in vocal brand personality perception. It replicates and extends the analysis presented in Essay 2, which empirically examined how brand personalities are perceived through voice in a German-speaking context. Following the German approach, the first objective of Essay 3 is determining which brand personalities can be perceived through voice in an American English-speaking context by developing an American BVP-Scale. Further, the study's second objective is developing an American BVP-Model that determines how female and male voices induce the perception of brand voice personality dimensions. The third objective is determining cultural differences in vocal brand personality perceptions between the United States and Germany. Therefore, a comparison of the German and American BVP-Scales and BVP-Models is made, highlighting the cultural differences and similarities.

Essay 3 is aligned with the German approach to ensure consistency in methodological and analytical procedures, facilitating a meaningful comparison of the findings. Therefore, the study employs a correlational design with external judgments by naïve listeners of brand personality perception through voice. In an online survey, American listeners were asked to evaluate two voices on the extent to which they associated the voice with the provided brand personality traits. The personality perception rating was based on the personality traits derived from the BPS of Aaker (1997), Geuens et al. (2009), and Grohmann (2009), which were also used in the German study. Further, voices from the Buckeye corpus of conversational speech were selected as suitable vocal stimuli for the study (Pitt et al., 2007). The survey yielded 1,958 individual voice ratings from 979 participants for the final data set (M_{age} = 34.35; 65.6% female).

Following the first research objective, the study first identifies which brand personalities can be perceived through voice alone in American English. To this end, analog to German research, the American BVP-Scale is developed based on a series of factor analyses and statistical procedures to explore the best model fit and establish important validities. The American BVP-Scale demonstrates to be a reliable and valid measurement tool, as evidenced by its strong indicator and construct reliability, internal consistency, discriminant and convergent validity. The developed scale represents a reflective first-order model with three American brand voice personality dimensions and eleven brand voice personality traits. The dimensions are labeled and defined as follows:

1) **Sincerity**: This dimension describes a brand voice personality perceived as honest, responsible, and sincere.
2) **Sensitivity**: This dimension describes a brand voice personality that is perceived as sensitive, tender, sentimental, and expresses tender feelings.
3) **Excitement**: This dimension describes a brand voice personality perceived as cool, adventurous, trendy, and exciting.

Following the second research objective, an American BVP-Model is developed by investigating which linear combinations of voice parameters induce the perceptions of brand voice personality dimensions. In doing so, a distinction is made between the perception of male and female voices. The model is constructed based on an SEM with a group code approach to estimate the speaker's gender-specific effects in which the perceptions of the three brand voice personality dimensions, sincerity, sensitivity, and excitement, are related to eleven acoustical measures.[2] In accordance with the American BVP-Model, voice profiles are derived for each brand voice personality dimension providing an overview of the combination of acoustic measures that is most favorable for the respective personality perception in female and male voices:

1) **Sincerity** in American brands is perceived when *female voices* are fluent in speaking, creaky, rough, and show no brightness; and *male voices* are breathy, bright, and demonstrate dynamic pitch variability and monotonic loudness variability when speaking.
2) **Sensitivity** in American brands is perceived when *female voices* are creaky, breathy, and without brightness but with a dynamic pitch intonation and a monotone loudness intonation; and *male voices* are breathy with monotone intonation and hesitant speech.
3) **Excitement** in American brands is perceived when *female voices* are breathy and creaky, have a less bright and clear articulation and a fluent speaking style; and *male voices* demonstrate a fluent and quick speaking style.

Following the third research objective, the American and German BVP-Scales and BVP-Models are compared to identify cross-cultural differences in the perception of brand personalities.

Comparison of American and German BVP-Scales: Both scales consist of the brand voice personality dimensions of sincerity, sensitivity, and excitement and encompass eleven brand voice personality traits. Therefore, similar brand personality dimensions are perceived through voice in both cultures. The most remarkable difference lies in the excitement

[2] speaking rate, average silent pause duration, silent pause frequency, articulation rate, intensity variability, mean and range of the fundamental frequency (f0), h-h2, spectral slope and tilt, and shimmer

dimensions, which must be considered when comparing the two personalities and their vocal perception. The item composition of the respective dimensions may be attributed to cultural differences in the perception of personalities in general. Additionally, differences could arise from the translations of the traits, which result in slight semantic variations that significantly impact the perception of voices.

Comparison of American and German BVP-Models: Despite the perception of similar brand personality dimensions through voice (sincerity, sensitivity, excitement), the cross-cultural comparison of the American and German BVP-Models and their derived brand voice profiles reveals that the profiles exhibit more differences than similarities in most cases. The female brand voice profiles of brands perceived as sincere and exciting have the most remarkable resemblance to one another. The voice profiles of sensitive brands are distinct in both cultures, as are the male voice profiles of sincere brands. Concerning exciting brands, the German male brand voice profile is the opposite of the American male profile. The findings support the theory that cross-cultural variations in learned social interaction norms and rules are crucial in (vocal) personality perception. Further, the study highlights the necessity for organizations with an international presence to select or design distinct brand voices depending on their customers' culture.

In conclusion, the value of the cross-cultural study of Essay 3 lies in its capacity to assess the generalizability of prior empirical findings and to contribute to a better understanding of the cultural influence on vocal brand personality perception. The findings assist companies operating in the U.S. and internationally in selecting and designing their brand voices to enhance the user experience and build customer trust, loyalty, and engagement with the brand (Ha & Stoel, 2009). In addition to the managerial recommendations, the findings contribute to the field of vocal personality perception and voice marketing research in three ways. First, they contribute to the limited research on the vocal perception of brand personalities, as previous research focused on examining human personalities. Second, the study of Essay 3 is the first to investigate the vocal brand personality perception within a cross-cultural context. Third, the study's character of replication contributes to the applicability of the utilized methods and models to facilitate the development of global strategies on vocal brand personality perception.

Synopsis References

Aaker, J. L. (1997). Dimensions of Brand Personality. *Journal of Marketing Research*, 34(3), 347–356. https://doi.org/10.1177/002224379703400304

Addington, D. W. (1968). The relationship of selected vocal characteristics to personality perception. *Speech Monographs*, 35(4), 492–503. https://doi.org/10.1080/03637756809375599

Aggarwal, P., & McGill, A. L. (2007). Is that car smiling at me? Schema congruity as a basis for evaluating anthropomorphized products. *Journal of Consumer Research*, 34(4), 468–479. https://doi.org/10.1086/518544

Allport, G. W., & Cantril, H. (1934). Judging personality from voice. *The Journal of Social Psychology*, 5(1), 37–55.

Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, 37(5), 715–727. https://doi.org/10.1037/0022-3514.37.5.715

Atkinson, R. K., Mayer, R. E., & Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology*, 30(1), 117–139. https://doi.org/10.1016/j.cedpsych.2004.07.001

Azoulay, A., & Kapferer, J.-N. (2003). Do brand personality scales really measure brand personality? *Journal of Brand Management*, 11(2), 143–155. https://doi.org/10.1057/palgrave.bm.2540162

Bank of America. (n.d.). *Erica is here for you, your life and your goals: Spend, save and plan smarter with Erica, your virtual financial assistant*. https://promotions.bankofamerica.com/digitalbanking/mobilebanking/erica

Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding voice perception. *British Journal of Psychology*, 102(4), 711–725. https://doi.org/10.1111/j.2044-8295.2011.02041.x

Bosak, O., Eisenbeiss, M., & Klein, K. (2023). Can you hear my personality? A conceptualization of a brand voice based on brand personality. In *52nd Proceedings of the European Marketing Academy (EMAC).*

Brown, B. L., Giles, H., & Thakerar, J. N. (1985). Speaker evaluations as a function of speech rate, accent and context. *Language & Communication*, 5(3), 207–220. https://doi.org/10.1016/0271-5309(85)90011-4

Caprara, G. V., Barbaranelli, C., & Guido, G. (2001). Brand personality: How to make the metaphor fit? *Journal of Economic Psychology*, 22(3), 377–395. https://doi.org/10.1016/s0167-4870(01)00039-3

Chandler, J., & Schwarz, N. (2010). Use does not wear ragged the fabric of friendship: Thinking of objects as alive makes people less willing to replace them. *Journal of Consumer Psychology,* 20(2), 138–145. https://doi.org/10.1016/j.jcps.2009.12.008

Cheng, A. (2018, June 20). Amazon-Marriott Deal Will Make Alexa A Hotel Butler, But The Implications Range Far Wider. *Forbes*. https://www.forbes.com/sites/andriacheng/2018/06/19/amazons-marriott-deal-is-way-beyond-alexa-as-your-new-hotel-butler/

Dahl, D. W. (2010). Understanding the role of spokesperson voice in broadcast advertising. In A. Krishna (Ed.), *Sensory marketing: Research on the sensuality of products* (pp. 169–182). Routledge. https://doi.org/10.4324/9780203892060-21

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, 69(3), 561–568. https://doi.org/10.1016/j.anbehav.2004.06.012

Geuens, M., Weijters, B., & Wulf, K. de (2009). A new measure of brand personality. *International Journal of Research in Marketing*, 26(2), 97–107. https://doi.org/10.1016/j.ijresmar.2008.12.002

Gobl, C., & Chasaide, A. N. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, 40(1-2), 189–212. https://doi.org/10.1016/S0167-6393(02)00082-1

Grohmann, B. (2009). Gender Dimensions of Brand Personality. *Journal of Marketing Research,* 46(1), 105–119. https://doi.org/10.1509/jmkr.46.1.105

Guha, A., Bressgott, T., Grewal, D., Mahr, D., Wetzels, M., & Schweiger, E. (2023). How artificiality and intelligence affect voice assistant evaluations. *Journal of the Academy of Marketing Science,* 51(4), 843–866. https://doi.org/10.1007/s11747-022-00874-7

Ha, S., & Stoel, L. (2009). Consumer e-shopping acceptance: Antecedents in a technology acceptance model. *Journal of Business Research,* 62(5), 565–571. https://doi.org/10.1016/j.jbusres.2008.06.016

Hörner, T. (2023). Marken auf Sprachassistenten. In T. Hörner (Ed.), *Marketing mit Sprachassistenten: So setzen Sie Alexa & Co strategisch erfolgreich ein* (2nd ed., pp. 191–222). Springer Gabler. https://doi.org/10.1007/978-3-658-40196-2_6

Jurafsky, D., & Martin, J. H. (2020). *Speech and language processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (3rd ed.).

Ketrow, S. M. (1990). Attributes of a telemarketer's voice and persuasiveness.: A review and synthesis of the literature. *Journal of Direct Marketing*, 4(3), 7–21. https://doi.org/10.1002/dir.4000040304

Klasmeyer, G., & Sendlmeier, W. F. (1997). The classification of different phonation types in emotional and neutral speech. *International Journal of Speech, Language and the Law*, 4(1), 104–124. https://doi.org/10.1558/ijsll.v4i1.104

Kramer, E. (1964). Personality Stereotypes in Voice: A Reconsideration of the Data. *The Journal of Social Psychology,* 62(2), 247–251. https://doi.org/10.1080/00224545.1964.9919523

Krauss, R. M., Freyberg, R., & Morsella, E. (2002). Inferring speakers' physical attributes from their voices. *Journal of Experimental Social Psychology*, 38(6), 618–625. https://doi.org/10.1016/S0022-1031(02)00510-3

Kreiman, J., & Sidtis, D. (2011). Perception of Emotion and Personality from Voice. In *Foundations of Voice Studies* (pp. 302–360). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444395068.ch9

Kreutzer, R. T., & Vousoghi, D. S. (2020). *Voice-Marketing: Der Siegeszug der digitalen Assistenten.* Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-29474-8_2

Kumar, A. (2018). Story of Aaker's brand personality scale criticism. *Spanish Journal of Marketing - ESIC*, 22(2), 203–230. https://doi.org/10.1108/SJME-03-2018-005

Kumar, V., Dixit, A., Javalgi, R. G., & Dass, M. (2016). Research framework, strategies, and applications of intelligent agent technologies (IATs) in marketing. *Journal of the Academy of Marketing Science*, 44(1), 24–45. https://doi.org/10.1007/s11747-015-0426-9

Lee, H., & Cho, C.-H. (2020). Uses and gratifications of smart speakers: Modelling the effectiveness of smart speaker advertising. *International Journal of Advertising*, 39(7), 1150–1171. https://doi.org/10.1080/02650487.2020.1765657

Lee, H. O., & Boster, F. J. (1992). Collectivism-Individualism in Perceptions of Speech Rate. *Journal of Cross-Cultural Psychology*, 23(3), 377–388. https://doi.org/10.1177/0022022192233008

Mari, A., Mandelli, A., & Algesheimer, R. (2020). The Evolution of Marketing in the Context of Voice Commerce: A Managerial Perspective. In F. F.-H. Nah & K. Siau (Eds.), *Lecture Notes in Computer Science. HCI in Business, Government and Organizations* (Vol. 12204, pp. 405–425). Springer International Publishing. https://doi.org/10.1007/978-3-030-50341-3_32

McAleer, P., & Belin, P. (2019). The Perception of Personality Traits from Voices. In S. Frühholz & P. Belin (Eds.), *The Oxford Handbook of Voice Perception* (1st ed., pp. 585–606). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198743187.013.26

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *PLOS ONE*, 9(3), 1–9. https://doi.org/10.1371/journal.pone.0090779.t001

Medical Xpress (2024, June 21). Our brains respond differently to human and AI-generated speech, but we still struggle to tell them apart. *Medical Xpress*. https://medicalxpress.com/news/2024-06-brains-differently-human-ai-generated.html

Mehrabian, A., & Ferris, S. R. (1967). Inference of attitudes from nonverbal communication in two channels. *Journal of Consulting Psychology*, 31(3), 248–252. https://doi.org/10.1037/h0024648

Melzner, J., Bonezzi, A., & Meyvis, T. (2022). EXPRESS: Information Disclosure in the Era of Voice Technology. *Journal of Marketing*, 87(4), 491-509. https://doi.org/10.1177/00222429221138286

Mercedes-Benz Group AG. (2022, April 19). *MBUX reaches a new level* [Press release]. https://group.mercedes-benz.com/innovation/digitalisation/connectivity/mbux-interior-assist.html

Moore, W. E. (1939). Personality Traits and Voice Quality Deficiencies. *Journal of Speech Disorders,* 4(1), 33–36. https://doi.org/10.1044/jshd.0401.33

Nam, J., Ekinci, Y., & Whyatt, G. (2011). Brand equity, brand loyalty and consumer satisfaction. *Annals of Tourism Research*, 38(3), 1009–1030. https://doi.org/10.1016/j.annals.2011.01.015

North, A. C., Mackenzie, L. C., Law, R. M., & Hargreaves, D. J. (2004). The Effects of Musical and Voice "Fit" on Responses to Advertisements. *Journal of Applied Social Psychology*, 34(8), 1675–1708. https://doi.org/10.1111/j.1559-1816.2004.tb02793.x

Oksenberg, L., Coleman, L., & Cannell, C. F. (1986). Interviewers' Voices and Refusal Rates in Telephone Surveys. *Public Opinion Quarterly*, 50(1), 97. https://doi.org/10.1086/268962

OpenAI. (2023). ChatGPT can now see, hear, and speak. *OpenAI*. https://openai.com/index/chatgpt-can-now-see-hear-and-speak/

OpenAI. (2024). Navigating the challenges and opportunities of synthetic voices. *OpenAI*. https://openai.com/index/navigating-the-challenges-and-opportunities-of-synthetic-voices/

Packard, G., & Berger, J. (2024). The Emergence and Evolution of Consumer Language Research. *Journal of Consumer Research*, 51(1), 42–51. https://doi.org/10.1093/jcr/ucad013

Pang, E. (2021). Exploring Cultural Awareness When Designing Voice UX: Are voice assistants culturally sensitive? *Medium*. https://medium.com/55-minutes/the-importance-of-cultural-awareness-in-design-14a5c0ae2ad8

Pear, T. H. (1931). *Voice and Personality* (1st ed.). Chapman and Hall, Ltd. https://doi.org/10.1038/128657a0

Peng, Y., Zebrowitz, L. A., & Lee, H. K. (1993). The Impact of Cultural Background and Cross-Cultural Experience on Impressions of American and Korean Male Speakers. *Journal of Cross-Cultural Psychology*, 24(2), 203–220. https://doi.org/10.1177/0022022193242005

Peterson, R. A., Cannito, M. P., & Brown, S. P. (1995). An Exploratory Investigation of Voice Characteristics and Selling Effectiveness. *Journal of Personal Selling & Sales Management*, 15(1), 1–15. https://doi.org/10.1080/08853134.1995.10754008

Pisanski, K., & Rendall, D. (2011). The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *The Journal of the Acoustical Society of America*, 129(4), 2201–2212. https://doi.org/10.1121/1.3552866

Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., & Fosler-Lussier, E. (2007). *Buckeye Corpus of Conversational Speech* (2nd release). Columbus, OH: Department of Psychology, Ohio State University (Distributor) [www.buckeyecorpus.osu.edu]

Rauschnabel, P. A., & Ahuvia, A. C. (2014). You're so lovable: Anthropomorphism and brand love. *Journal of Brand Management*, 21, 372–395. https://doi.org/10.1057/bm.2014.14

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people*. Cambridge University Press.

Rodero, E. (2013). The Perception of a Broadcasting Voice. *US-China Education Review*, 3(4), 225–230.

Rodero, E. (2017). Effectiveness, attention, and recall of human and artificial voices in an advertising story. Prosody influence and functions of voices. *Computers in Human Behavior,* 77, 336–346. https://doi.org/10.1016/j.chb.2017.08.044

Scherer, K. R. (1972). Judging personality from voice: A cross-cultural approach to an old issue in interpersonal perception. *Journal of Personality*, 40(2), 191–210. https://doi.org/10.1111/j.1467-6494.1972.tb00998.x

Schwartz, E. H. (2020). Australians Can Ask Alexa for a Personalized Bottle of Coke. *voicebot.ai*. https://voicebot.ai/2020/01/30/australians-can-ask-alexa-for-a-personalized-bottle-of-coke/

Schweinberger, S. R., & Zäske, R. (2019). Perceiving Speaker Identity from the Voice. In S. Frühholz & P. Belin (Eds.), *The Oxford Handbook of Voice Perception* (1st ed., pp. 538–560). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198743187.013.24

Sharf, D. J., & Lehman, M. E. (1984). Relationship between the speech characteristics and effectiveness of telephone interviewers. *Journal of Phonetics*, 12(3), 219–228. https://doi.org/10.1016/S0095-4470(19)30878-2

Skjegstad, C. L., & Frühholz, S. (2024). Neural dynamics of processing natural and digital emotions [Poster Session]. In *FENS 2024 Conference Programm*. https://fens2024.abstractserver.com/program/#/details/presentations/4774

Steiner, P. (2018). *Sound Branding: Grundlagen akustischer Markenführung* (3rd ed.). Springer Fachmedien Wiesbaden. https://link.springer.com/content/pdf/10.1007%2F978-3-658-22638-1.pdf https://doi.org/10.1007/978-3-658-22638-1

van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, 38(3), 253–265. https://doi.org/10.1177/002383099503800303

VML. (2023). *A built-in Alexa skill that rewards viewers every time a Coke pops on screen*. https://www.vml.com/work/i-see-coke

Volkswagen Group. (2024). *ChatGPT is now available in many Volkswagen models*. https://www.volkswagen-group.com/en/articles/chatgpt-is-now-available-in-many-volkswagen-models-18464

Waaramaa, T., Lukkarila, P., Järvinen, K., Geneid, A., & Laukkanen, A.-M. (2021). Impressions of Personality from Intentional Voice Quality in Arabic-Speaking and Native Finnish-Speaking Listeners. *Journal of Voice: Official Journal of the Voice Foundation,* 35(2), 326.e21-326.e28. https://doi.org/10.1016/j.jvoice.2019.09.003

Wagner, K., & Schramm-Klein, H. (2019). Alexa, Are You Human? Investigating Anthropomorphism of Digital Voice Assistants - A Qualitative Approach. In *Proceedings of the Fortieth International Conference on Information Systems (ICIS), Munich 2019.*

Wee, T. T. T. (2004). Extending human personality to brands: The stability factor. *Journal of Brand Management*, 11(4), 317–330. https://doi.org/10.1057/palgrave.bm.2540176

Wenden, A. (2020). Alexa Speaking Styles and Emotions Now Available in Additional Languages. *Amazon*. https://developer.amazon.com/en-US/blogs/alexa/alexa-skills-kit/2020/11/alexa-speaking-styles-emotions-now-available-additional-languages

Wiener, H. J. D., & Chartrand, T. L. (2014). The Effect of Voice Quality on Ad Efficacy. *Psychology & Marketing*, 31(7), 509–517. https://doi.org/10.1002/mar.20712

Wu, Q., Liu, Y., Li, D., Leng, H., Iqbal, Z., & Jiang, Z. (2021). Understanding one's character through the voice: Dimensions of personality perception from Chinese greeting word "Ni Hao". *The Journal of Social Psychology*, 161(6), 653–663. https://doi.org/10.1080/00224545.2020.1856026

Zäske, R., Skuk, V. G., Golle, J., & Schweinberger, S. R. (2020). The Jena Speaker Set (JESS) - A database of voice stimuli from unfamiliar young and old adult speakers. *Behavior Research Methods,* 52(3), 990–1007. https://doi.org/10.3758/s13428-019-01296-0

Zoghaib, A. (2017). The contribution of a brand spokesperson's voice to consumer-based brand equity. *Journal of Product & Brand Management*, 26(5), 492–502. https://doi.org/10.1108/JPBM-06-2016-1230

Essay 1

Voices that Resonate: A Systematic Review and Future Research Agenda of Vocal Personality Perception for Brand Voice Selection

Author: Olga Bosak

Abstract

Integrating voice touchpoints into an organization's branding strategy facilitates the creation of personalized customer experiences that are convenient, rapid, and easily accessible. In addition, through voice-based devices, consumer-brand interaction is increasingly natural and social as brands can "talk back" by engaging in two-way communication with their consumers. However, the advent of voice technology challenges marketers regarding their brand's communication, as voice-based devices typically lack visual content. Consequently, brands must rely exclusively on auditory communication, making the voice the sole medium through which the brand's personality is conveyed. Despite the existing body of research on personality perception through voice, a summary of the findings is lacking. The question for marketers of which voice characteristics are essential when selecting a brand voice to reflect the desired personality remains. This study addresses this issue by presenting the first comprehensive review of personality perception through voice parameters, encompassing nearly 85 years of published research. Additionally, it outlines how findings on the vocal perception of human personality can be applied to brand personalities, along with managerial recommendations for brand voice selection. In light of the identified research gaps, this study proposes a research agenda for the future.

Keywords: brand voice, brand personality, personality perception, voice perception, voice AI, voice marketing, systematic literature review, future research agenda

1. Introduction

The use of conversational agents, voice assistants, social robotics, and other voice-based artificial intelligence (AI)[1] technologies has grown exponentially in recent years (voicebot.ai & Business Wire, 2020). As a result, disembodied voices have become an integral part of daily life worldwide. Organizations have begun to utilize these emerging voice touchpoints to provide personalized customer experiences that are convenient, fast, and easy to access (Kreutzer & Vousoghi, 2020; Mari et al., 2020; Steiner, 2018). For example, automobile manufacturers such as BMW and Mercedes-Benz have integrated voice-controlled technologies into their vehicles that are capable of remembering and recalling the driver's preferences or preferred routes (BMW Group, 2018; Mercedes-Benz Group AG, 2022). Additionally, virtual assistants, such as "Erica" from Bank of America, are utilized in the financial and banking sector to facilitate customer financial management through personalized voice interactions (Bank of America, n.d.). Other companies have adopted third-party voice assistants, such as Alexa or Google Home, rather than developing their own. Marriott is one such company that has integrated Alexa into its rooms, utilizing it as a "hotel butler" (Cheng, 2018) to enhance the guest experience.

These interactions between consumers and voice assistants demonstrate a new form of consumer-brand communication. Traditionally, brand voice interactions with customers have been one-way or based on predefined Q&A prompts, such as automated phone interventions through interactive voice response systems (IVR; Inam et al., 2017; Jurafsky & Martin, 2020). Now, brands can "talk back" (Packard & Berger, 2024, p. 46) by engaging in two-way, complex, and increasingly human-like interactions with consumers through the use of voice technology (Jurafsky & Martin, 2020; V. Kumar et al., 2016). The advancement of social and

[1] **artificial intelligence (AI)** = "Systems that display intelligent behavior by analyzing their environment and taking actions – with some degree of autonomy – to achieve specific goals." Sheikh et al. (2023, p. 20); **voice/ voice-based AI** = AI technologies enabled to perform speech recognition and natural language processing (NLP) to engage in natural dialogues with users (Wang et al., 2023)

natural communication is enabled not only by the growing realism of the assistants' voices (Medical Xpress, 2024; Skjegstad & Frühholz, 2024), but also by the tendency of individuals to ascribe human characteristics to computers and devices (Reeves & Nass, 1996). This anthropomorphization of voice assistants and brands offers many benefits that companies can leverage, as it positively influences product evaluations (Aggarwal & McGill, 2007), enhances product likeability (Chandler & Schwarz, 2010), and boosts brand recall, affection, and loyalty (Rauschnabel & Ahuvia, 2014).

However, the rise of voice technology presents a new challenge for marketers regarding their brand's conversation with customers, as voice devices and platforms usually lack visual content. As a result, brands must rely solely on auditory communication, making the voice the primary focus of the exchange. Further, as voice assistants predominantly use the same voice for all interactions, such as the voice of Alexa or Siri, there is a risk that customers will not connect the voice content to a specific brand (Paluch & Wittkop, 2020). Therefore, choosing an appropriate and unique voice representing an organization and its brand with a preferred attitude and personality is essential in improving the overall customer voice experience and differentiating from the competition (McAleer & Belin, 2019). But what is the appropriate voice to reflect a desired brand personality? Building on research findings into the vocal perception of personalities, this study is designed to answer this question.

Research on personality perception through voice has been conducted for nearly a century, resulting in the development of numerous theories and the generation of empirical results (Allport & Cantril, 1934; Moore, 1939; Pear, 1931). Consequently, it has increasingly been utilized in marketing and consumer research by investigating the role of spokesperson voices in telemarketing, direct selling, radio, or TV (Dahl, 2010; Ketrow, 1990). Despite the substantial corpus of empirical evidence on vocal personality perception and its marketing implications, marketers face challenges in selecting the appropriate brand voice for voice-based devices

(Kreutzer & Vousoghi, 2020). There is a lack of a comprehensive summary, overview, or guide specifying which voice characteristics lead to personality perceptions. Therefore, marketers predominantly rely on their gut feelings or experiences to select a brand voice (Dahl, 2010; Wiener & Chartrand, 2014). In addition, perceptual studies have mainly focused on the vocal perception of human personalities, not brand personalities.

Brand personalities are conceptualized as brands possessing humanlike traits and were found to function the same way human personalities do (Wee, 2004). However, research on the compatibility between human and brand personalities has also revealed that not all personality traits that describe a human can be transferred to brands (A. Kumar, 2018). One possible reason for this discrepancy is that personality traits applied to humans may have different meanings when applied to brands (Caprara et al., 2001). Also, some human personality traits do not apply to brands; for instance, human traits like “neurotic fatigue” (Azoulay & Kapferer, 2003, p. 149). Further, brands are more likely to be associated with excessive characteristics such as “sophisticated” or “rugged,” which humans desire yet do not possess (Aaker, 1997). Consequently, while the concepts of human and brand personalities resemble each other, a distinction must be made between which human personality traits apply to brands. Therefore, research on the vocal perception of relevant personalities applicable to brands is needed.

This study addresses this issue by identifying, summarizing, and classifying existing research on vocal human personality perception through a systematic literature review. Further, based on the empirical findings on human personalities, conclusions are drawn on the perception of brand personalities. In doing so, the study follows the first developed and well-known brand dimensions of Aaker (1997), sincerity, excitement, competence, sophistication, and ruggedness, and derives vocal profiles for these personality dimensions. In addition to this primary research objective, the literature review is guided by further research questions concerning the stimulus material (speaker’s gender, language, speech data type) and methodology (study design, assessment method) of perceptual studies. This study examines how the literature on vocal personality perception addresses these aspects to formulate managerial guidance for determining the appropriate voice to fit a brand personality. Further, the findings and identified research gaps are used to derive a future research agenda for vocal brand and human personality perception.

This study contributes to the literature on vocal personality perception and voice marketing in three ways. First, it represents the first comprehensive systematic review of human personality perception through voice parameters, covering nearly 85 years of published research on the topic with 52 articles. Second, it shows how findings on vocal perception of human personality can be applied to brand personalities by deriving vocal profiles for sincere, exciting, competent, and sophisticated brands to guide managers in selecting brand voices. Third, it outlines five avenues for future research based on current advances, such as developments in speech processing and analysis technology and the changing role of genders in society.

The remaining sections are organized as follows: First, background information is provided on voice, including inferences drawn from voices, perception of personality through the voice, and the definition of brand personalities. Building on this theory, the research questions are formulated, and the derived research framework of this systematic literature review is demonstrated. Next, the methodology of a systematic literature review is described in detail, and the resulting findings are presented on a macro and micro level. Finally, the study concludes with a proposal for a future research agenda, a summary of the results, and the limitations.

2. Theoretical Background

2.1. Inferences from Voice

Voice is produced when air from the lungs passes through the larynx, causing the vocal folds to vibrate (Clark, 2007; Pompino-Marschall, 2009; see Appendix A for a schematic diagram of speech production organs and the vocal tract). These vibrations are then modified by the articulators in the vocal tract, such as teeth, tongue, oral, and nasal cavity, producing signals in various frequencies and amplitudes that are captured in a soundwave (Frühholz & Belin, 2019; Klasmeyer & Sendlmeier, 1997). These soundwaves are registered acoustically and perceived as speech by the listener (Hildebrand et al., 2020).

Human speech generally provides information from at least two sources: a verbal channel, which encodes the semantic content of a message, and a paraverbal, i.e., vocal, channel, which conveys paralinguistic information through variations in pitch, speech rate, loudness, and other voice parameters (Apple et al., 1979). The paraverbal channel can be used to perceive a variety of information about the speaker that goes beyond the content of a message (McAleer & Belin, 2019). Even without visual cues, such as hearing a voice on the telephone or radio, people can make inferences about a person based on paralinguistic information alone (Koutsombogera et al., 2020; Stern et al., 2021).

The characteristics attributed to a person by their voice can be divided into three categories: 1) physical markers, which refer to characteristics such as gender, age, height and weight, or health status (Afshan et al., 2018; Collins, 2000; Krahé & Papakonstantinou, 2020; Skoog Waller & Eriksson, 2016); 2) social markers indicating characteristics such as regional origin, educational status, and social class (Krauss et al., 2002; Stewart & Ryan, 1982); and 3) psychological markers providing characteristics of affective state, identity, and personality (Klasmeyer & Sendlmeier, 1997; Zäske et al., 2020).

In particular, the study of psychological markers has been the focus of research in recent decades, with the identification and recognition of affective state from voice being the most investigated (Krauss et al., 2002). Consequently, emotional arousal is the most easily recognizable aspect of vocal communication. For example, by combining a slow speech rate, a high pitch, and a low loudness, sadness is perceived; by using a high pitch and pitch variability and a fast speech rate, happiness is perceived (Brave & Nass, 2009; Hildebrand et al., 2020).

2.2. Voice and Personality

In addition to the affective state as a psychological marker, acoustic information consists of cues about the speaker's personality, which is the most critical aspect of human differences (Mairesse et al., 2007). Emotions and personality are related concepts, with emotions being a temporary state and personality being a stable trait (Kreiman & Sidtis, 2011). That means individuals develop a personality early in life and tend to follow their personal preferences throughout life (Caspi & Roberts, 2001). In this context, personality research assumes that people differ in some general dimensions and that these differences shape a person's personality (van Mersbergen, 2011).

According to research, the impression of personality traits such as competence, dominance, and extraversion is significantly influenced by the paraverbal content of speech (Kreiman & Sidtis, 2011; Mehrabian & Ferris, 1967). As mentioned earlier, paraverbal content refers to acoustic parameters of the voice surrounding the semantic content of a message, such as speech rate, pitch, or loudness, which are expressed differently in each voice (Ketrow, 1990). All humans have a similar structure of the vocal tract system. However, differences in the length of vocal folds and slight variations in the acoustic parameters around a mean determine voice uniqueness, which leaves every person with an individual vocal signature, timbre, or voice color (Belin et al., 2011; Rodero, 2013).

Although voices are unique, psychologists have discovered correlations between several acoustic parameters and specific personality traits. Voice stereotypes are said to exist, meaning that single acoustic parameters or combinations of them cause the same personality to be perceived by listeners of different voices (Addington, 1968; Kramer, 1964). This phenomenon could be attributed to the fact that only a few parameters of a voice, namely pitch, loudness, and timing features, have been identified as the most critical parameters for perceiving a person's personality (D. Y. Huang et al., 2009; Imhof, 2010).

First-impression personality judgments are made subconsciously and can be established after just a few seconds (Tsantani et al., 2016). In their study, McAleer et al. (2014) showed that hearing a single word ("hello" in Scottish English) spoken by different voices led to personality perceptions by naïve listeners. Personality judgments influence social interactions, mate choice (Collins, 2000), voting behavior (Tigue et al., 2012), career opportunities (Mayew et al., 2013), and consumer choices (Lowe & Haws, 2017).

2.3. Brand Personality

Brand personality suggests that consumers perceive a brand as having a distinct personality representing characteristics and values. According to Aaker (1997), brand personality is defined as "the set of human characteristics associated with a brand" (p. 347). Azoulay and Kapferer (2003) extended this definition to refer to brand personality as "the set of human personality characteristics applicable and relevant to a brand" (p. 151) because not all human personality traits were found to be transferable to brands (Caprara et al., 2001; A. Kumar, 2018). The Theory of Animism, developed by Gilmore (1919), provides the basis for transferring human characteristics to a brand. This theory suggests that humans inherently need to animate non-living objects with human traits. The reasons for anthropomorphizing objects include the desire for familiarity, risk reduction, and simplifying the relationship with objects (Freling & Forbes, 2005).

Although the concept of brand personality was coined in the 1950s, it was not until 1997 that Aaker developed a scale to measure brand personality. Aaker's (1997) brand personality scale (BPS) consists of five dimensions, sincerity, excitement, competence, sophistication, and ruggedness, and incorporates 15 brand personality facets and 42 traits (see **Figure 1**). These brand personalities are based on research results from the psychology of human personality, which generally agree on the "Big Five" personality dimensions, namely openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (Aaker, 1997; Bosnjak et al., 2007). The Big Five dimensions are a product of studying the natural language terms people use to describe themselves and others (Goldberg, 1993). This model's widespread dissemination and use is also why the Big Five personalities are predominantly utilized in linguistic personality research. Since human personalities were frequently used as a foundation for the development of BPS, they often correlate with the Big Five at the item or dimension level. For example, Aaker's (1997) dimensions of sincerity, excitement, and competence correlate with the Big Five dimensions of agreeableness, extraversion, and conscientiousness, respectively (Aaker et al., 2001).

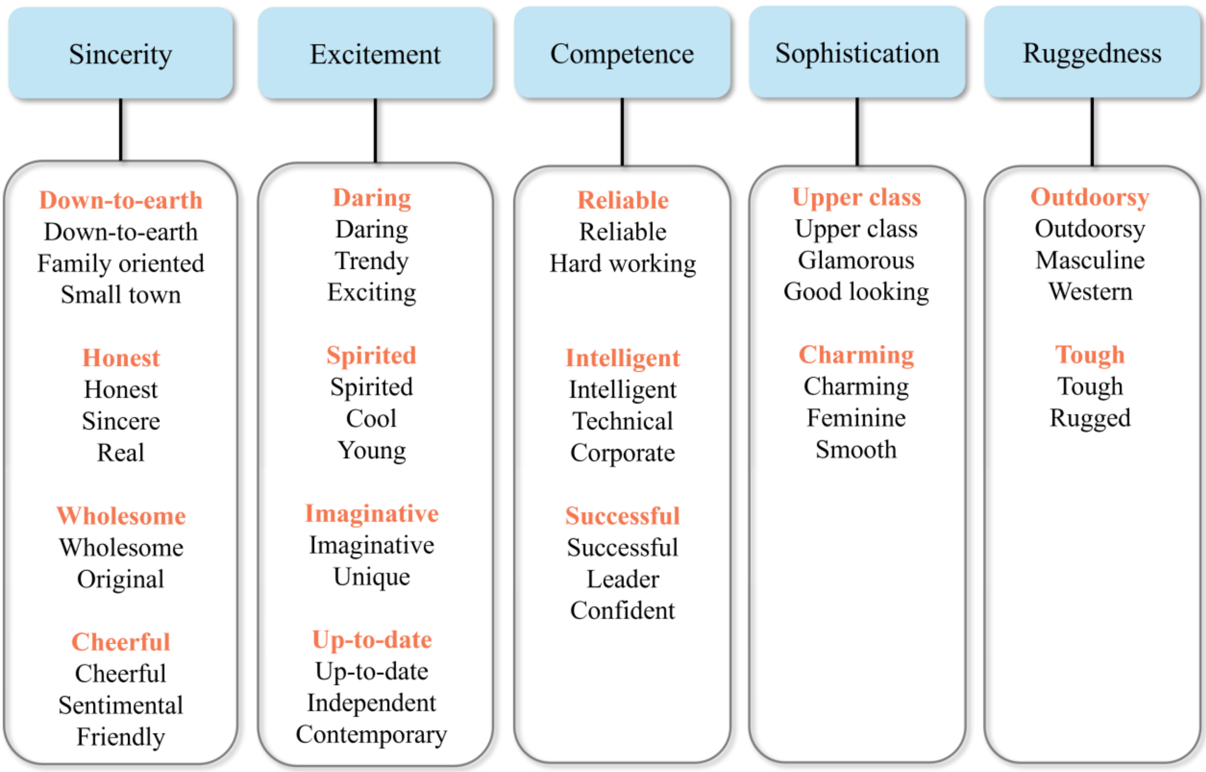


Figure 1. Brand Personality Framework by Aaker (1997)
Notes: Brand personality dimensions are at the top; facets are bold and orange; traits are thin and black.

3. Research Objective and Framework

Research on personality perception through voice has existed for nearly a century, producing numerous empirical results, concepts, and theories (Allport & Cantril, 1934; Moore, 1939; Pear, 1931). The emergence of innovative digital technologies and software designed to enhance speech processing analysis, along with advancements in phonetics and psychology, led research on voice perception to receive considerable attention between the 1960s and 1980s (Addington, 1968; Apple et al., 1979; Brown et al., 1985; Scherer, 1972). Nonetheless, despite the existing body of research on human personality perception through voice, there is a lack of research on vocal perceptions of personalities that are relevant and applicable to brands. An initial search for related literature in the primary scientific databases, such as Scopus or Web of Science, yielded no relevant results. A further examination of Lens, a research platform that provides analytical tools for scholarly literature, has revealed the existence of 60 scholarly works on brand personality/identity in combination with voice since 1993[2] (see **Figure 2**; The Lens, 2023). On closer examination, however, it became apparent that all identified studies had a different research objective regarding brand voice perception. None of them investigated how or which brand personalities are perceived through voice.

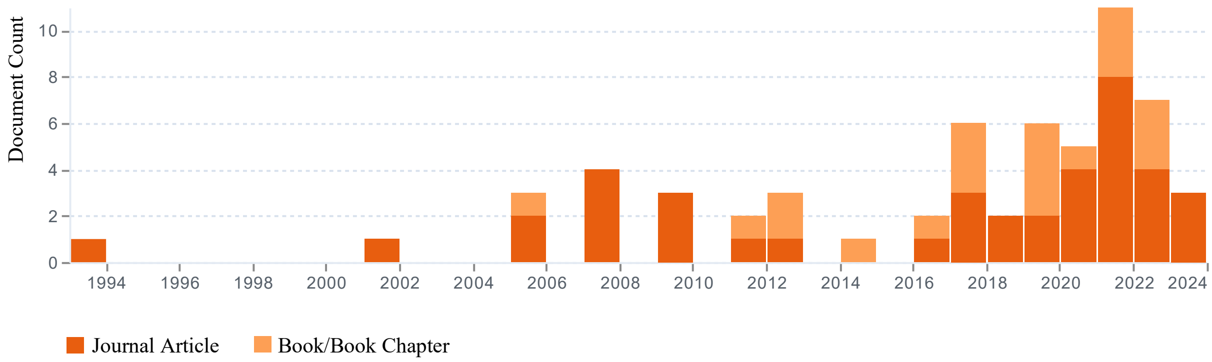


Figure 2. Scholarly Work on Brand Personality Perception Through Voice Over Time
Notes: Stacked bar chart; N = 60. **Source:** The Lens (2023).

In response to the identified gap in research on brand personality perception through voice, this study is the first to identify, summarize, and classify existing research on human personality perception. Based on the empirical findings on the vocal perception of human personalities, inferences are made on the vocal perception of brand personalities. In doing so, the

[2] search command: title:("brand personality" AND voice) OR title:("brand identity" AND voice) OR title:(brand OR voice)

study follows the brand personality dimensions of the BPS of Aaker (1997), as this scale has gained extensive use and acceptance in marketing research due to its high reliability and validity. By providing a comprehensive summary of which voice parameters induce personality perceptions, this study aims to guide marketers in selecting the appropriate voice that fits their brand personality. Further, the findings and identified research gaps in the literature are used to derive a future research agenda for vocal brand and human personality perception. A systematic literature review on vocal personality perception research is undertaken to achieve these objectives.

3.1. Research Questions

Addressing the research objective of reviewing existing research on human personality perception through voice parameters, this study focuses on extracting empirical data on identified significant linear or non-linear relationships between voice parameters as independent variables and personalities as dependent variables. Thereby, it shall also be distinguished if correlations between personality traits and individual voice parameters (univariate analysis) or combinations of voice parameters (multivariate analysis) were investigated. Thus, the study's primary research question (RQ) is:

RQ1: Which (combinations of) voice parameters induce human personality perceptions?

As perceptual studies differ in the stimulus material used and methodical approach, additional research questions guide the systematic review to assess the comparability of the findings.

Concerning the stimulus material, outcomes may differ based on the speaker's gender, the language, and the type of speech data. Research has shown that personalities can be perceived differently based on the speaker's gender due to socially constructed norms and values, influencing how they speak and, thus, how their personalities are perceived (Trouvain et al., 2021). In addition, culture influences how people use their voice, affecting how personalities are perceived through the voice (Krauss et al., 2002; Waaramaa et al., 2021). Further, it is crucial to distinguish between the speech data types utilized in perceptual studies, as personality traits are best reflected when speakers speak spontaneously and have no behavioral constraints (Johnstone & Scherer, 2000). Speech data can either be single vowels or sentences that are read, freely spoken (spontaneous speech), or freely spoken based on a guide or visual aid, e.g., a prepared speech (semi-spontaneous speech). Thus, within the review of the literature, the following research questions need to be answered concerning the stimulus material used:

RQ2a: What is the speaker's gender in the vocal perception of personality?

RQ2b: What language is used for the vocal perception of personality?

RQ2c: What type of speech data is used in the vocal perception of personality?

Concerning the methodology, phonetic experiments can be designed as correlation, cue manipulation, or cue synthesis studies. In a correlation study design, variables are measured without being manipulated. In experimental study design, variables are manipulated; voices can be manipulated subjectively through speakers, i.e., cue manipulation, or software, i.e., cue synthesis. In the case of cue manipulation, there is no assurance that the speaker is not manipulating other vocal parameters simultaneously. Cue synthesis offers a higher experimental control because single voice parameters can be manipulated while others remain unchanged. (Brown et al., 1985; Riding et al., 2006)

Further, research indicates that personality perception may differ in accuracy and validity based on the assessment method. Personality perceptions can be assessed through external judgments of listeners or self-ratings of the speaker. Due to values, experiences, and self-perceptions that are not apparent to external individuals, self-ratings on personality perception may lead to different outcomes than external judgments (Koutsombogera et al., 2020). Thus, within the review of the literature, the following research questions need to be answered concerning the methodological approach:

RQ3a: What study design is applied to the vocal perception of personality?

RQ3b: What assessment method is applied for the vocal perception of personality?

Following the research questions, this study suggests the research framework, as shown in **Figure 3**.

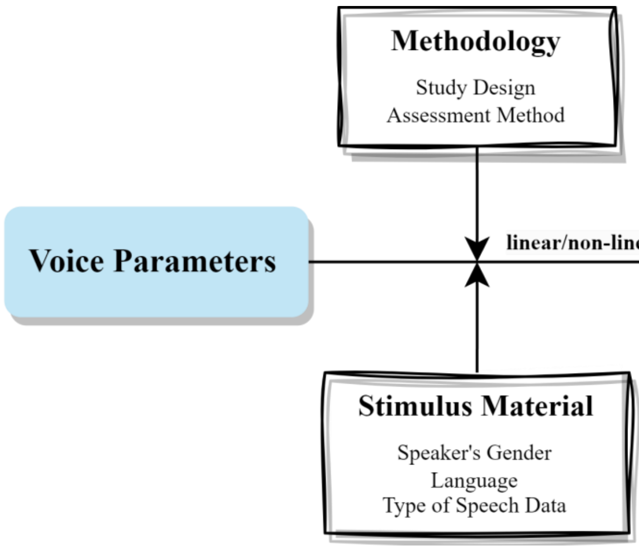


Figure 3. Research Framework of Literature Review on Vocal Personality Perception

4. Method

This study adopted a systematic literature review approach to identify, summarize, and classify existing research on personality perception through voice parameters. A systematic review describes a review of research results of clearly formulated research questions or objectives through a replicable, scientific, and transparent method (Kitchenham, 2004; Moher et al., 2010). This methodology ensures a comprehensive overview of research on a particular topic using a systematic search strategy according to predefined inclusion and exclusion criteria (Linnenluecke et al., 2020).

Literature on the perception of personality from voice can be found in language and communication sciences, linguistics, phonetics, psycholinguistics, psychology, consumer behavior, economics, and information systems. Due to the interdisciplinary character of the research topic, literature was not searched in selected journals but in scientific databases, which provided access to various online citations and literature databases. The electronic databases used for this literature search were Scopus, Web of Science, ProQuest, and PsycNet.

The search was performed in April 2023 within the title to find potentially relevant articles using the key terms *personality, perception,* and *voice* and their synonyms – *identity* for personality, *inference* and *judgment* for perception, and *vocal, acoustic,* and *speech* for voice – as search terms. These terms were used in all possible combinations with the Boolean operator AND to ensure the search was broad enough to find relevant literature. As the initial search within the title generated more than 10,000 articles, additional search terms as exclusion criteria were added with the Boolean operator (AND) NOT to reduce the results to an acceptable level (see Appendix B). In addition to excluding search terms, filters were set in the advanced search options to limit the results to peer-reviewed articles in English (see Appendix C).

4.1. Study Selection

A total of 3,026 articles were found through the described search process. Following the guidelines of the PRISMA statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) for systematic reviews (Moher et al., 2010), duplicates were removed, and articles were reviewed step by step according to the inclusion and exclusion criteria described in the following subsection. This process was done by screening first the title, then the abstract, and finally the full text. Finally, the reference lists of the fully read articles were reviewed to check whether potential articles were missed through the initial search on the databases, i.e., backward search.

4.1.1. Inclusion and Exclusion Criteria

To specifically address the research objectives and identify literature that was eligible for the scope of this study, the selected articles for the review needed to meet the following criteria: 1) they should include quantitative empirical findings in which significant correlations between voice parameters and human personality traits were found; 2) they should be published in a peer-reviewed journal; 3) they should be written in English.

Articles were only included if they met these three criteria. Additionally, articles were excluded if they

1) examined the influence of voice on personality as a whole and did not separate individual voice parameters, as in such studies, the impact of individual parameters could not be comprehended;
2) used visual stimuli for personality perception in combination with voice, as in such studies, it would be impossible to comprehend the specific role of the voice in personality perception;
3) used machine learning to identify (automatic) voice perception. In such studies, models are trained to recognize personalities in speech. The underlying concept of which voice parameters influence personality perception would not be comprehensible;
4) referred to a medical context, as such studies tend to examine the impact of therapies and treatments on the voice or how diseases/disorders such as Parkinson's, autism, or schizophrenia influence speech production and perception;
5) used non-native speakers or listeners, as in such studies, the influence of the listener's culture on personality perception could not be considered in isolation;
6) used children or older adults as speakers or listeners, as such studies tend to focus on the particularities of age in speech production and perception, which would be beyond the scope of this study.

In compliance with the inclusion/ exclusion criteria, the study selection process yielded 52 articles for the review. For a detailed overview of the search and selection process, see the

PRISMA flow diagram in **Figure 4**, which shows the number of studies included and excluded in each stage (Moher et al., 2010).

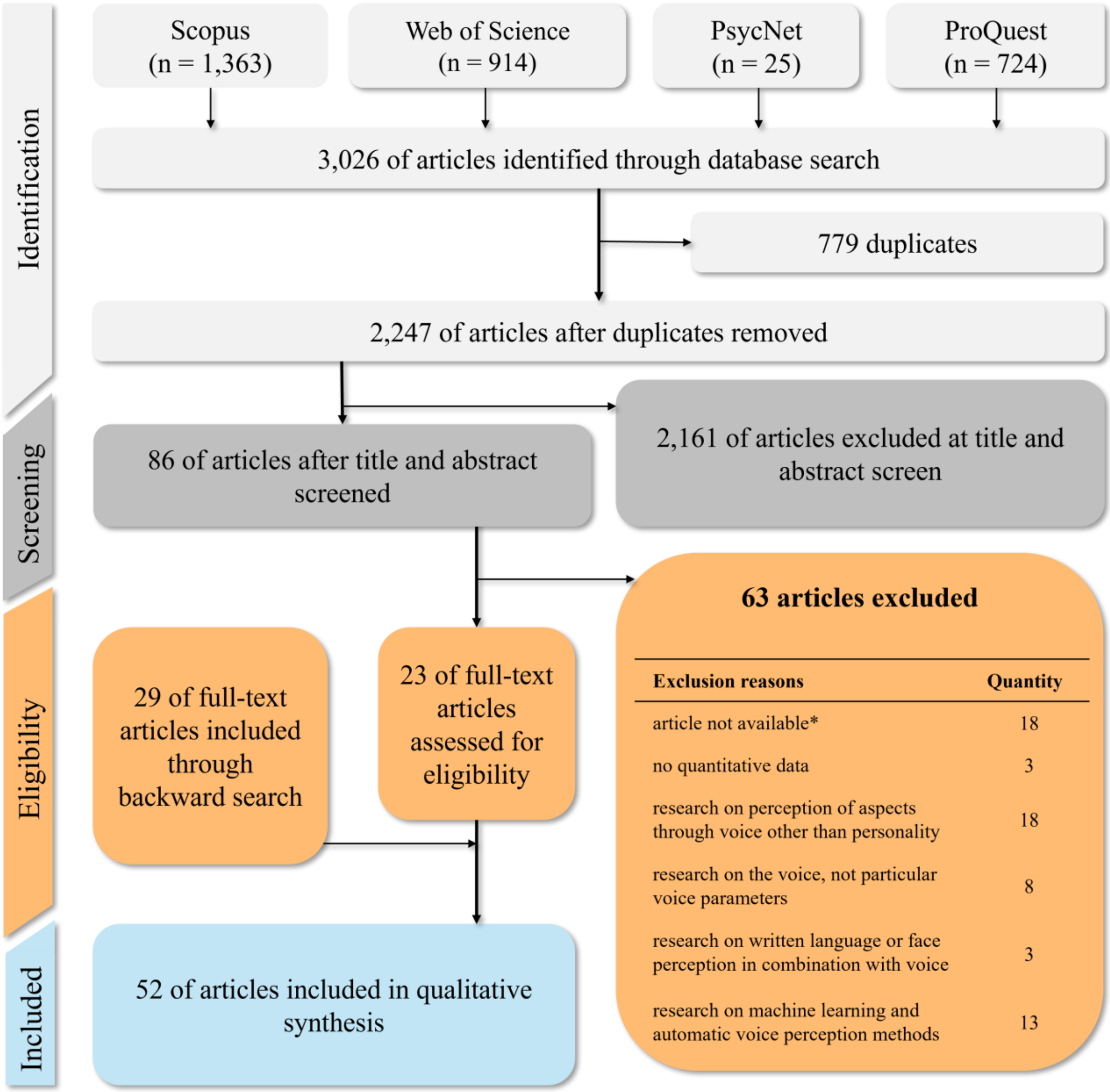


Figure 4. PRISMA Flow Diagram of the Search and Selection Process

Notes: *These articles were not accessible because they were older than 40 years. Therefore, contacting the authors was impossible as no current contact details were available.

4.2. Data Extraction

As recommended by Webster and Watson (2002), a concept matrix was compiled to extract data from the identified articles. Following the research questions and framework, the essential data to extract was categorized according to findings on 1) which (combinations of) voice parameters induced human personality perceptions, 2) used stimulus material, and 3) methodological aspects (see **Table 1**). Based on this categorization, the empirical findings of the studies were synthesized for the following analyses and interpretations.

Table 1. Concept Matrix of Extracted Data

Category	Extracted Data	Extracted Data Specifications
Vocal Personality Perception	Voice Parameters	
	Personalities	
	Relationship & Analysis	linear, non-linear; univariate, multivariate
Stimulus Material	Gender of Speaker	female, male
	Language	
	Type of Speech Data	read text, semi-spontaneous speech, spontaneous speech
Methodology	Study Design	correlation study, cue manipulation, cue synthesis
	Assessment Method of Personalities	self-ratings, external judgments

Appendices D-H provide a comprehensive overview of the 52 articles selected for the literature review, categorized according to the concept matrix and to which the following macro- and micro-level analyses refer.

5. Macro-Level Analysis

The macro-level analysis describes the identified articles and results on vocal personality perception on a descriptive level. First, it is explained in which journals the results were published, when research in vocal personality perception began, and how it has evolved. The extracted data is then summarized according to the categories of the concept matrix.

5.1. Categorization and Evolution of Research on Vocal Personality Perception

The extracted data comprises 52 articles from 31 journals, classified into two main research areas: 1) social, experimental, and cognitive psychology, and 2) acoustics, linguistics, languages, and communication (see Appendix I). This distribution is not unexpected, given that the perception of personalities affects the psyche and behavior of humans, while the voice and its parameters affect human communication and language. Therefore, most perceptual studies are found in interdisciplinary journals like *The Journal of the Acoustical Society of America*, *Animal Bevahiour*, *Communication Monographs*, and *PLOS ONE*, in which research contributes to understanding human communication and social psychology. Few studies have been conducted in research areas such as computer science or marketing. Only one study was published in the *Journal of Product and Brand Management*, which had a marketing focus by investigating the contribution of a brand spokesperson's voice to consumer-based brand equity (Zoghaib, 2017). This observation serves to reinforce the fact that there is a lack of research on the vocal perception of brand personalities.

Research on personality perception through voice has been found within the last 84 years. This literature is clustered into four periods, highlighting the advancements in phonetic technology and the evolving interests within the vocal personality perception research (see **Figure 5**). The *initial period* of research on vocal personality perception (1939-1970) encompasses three perceptual studies published by Moore (1939), Mallory and Miller (1958), and Addington (1968). In their original research, the investigators examined the relationship between individual voice characteristics and personality traits, establishing a foundation for subsequent research. In these studies, voice parameters could not be acoustically measured or altered but were subjectively classified by trained judges like speech experts or phoneticians. Additionally, Moore (1939) and Mallory and Miller (1958) examined personality self-ratings and focused on the Big Five personality dimensions like dominance, extraversion, or neuroticism. Addington (1968) was the first to conduct a cue manipulation study in which two speakers subjectively adjusted their voices to produce different voice qualities like nasality, breathiness, or roughness. Moreover, Addington (1968) was also the first to investigate personality traits other than the Big Five, such as carefulness, emotionality, activity, modesty, and politeness.

The *second period* (1971 - 1980) was mainly driven by the work of Brown, Smith, Strong, and Rencher (1972 - 1975). During that period, speech processing and analysis technology evolved, enabling researchers to undertake more complex vocal studies. In 1972, Brown et al. were the first to use technological devices to measure and alter speakers' pitch and speaking rate. Therefore, this period was dominated by cue synthesis studies and external judgments for assessing personality traits.

The *third period* (1981 – 1999) represents a time of perceptual studies in languages other than American English, with Brown et al. (1985) being the first to utilize British English speakers. Moreover, the early 1990s were dominated by cross-cultural studies in which the vocal perception of personalities was compared between Americans and Koreans (H. O. Lee & Boster, 1992; Peng et al., 1993) and between Dutch and Japanese (van Bezooijen, 1995). Additionally, most studies focused on investigating the effect of speech rate on the perception of personality (e.g., S. M. Smith & Shaffer, 1991; Street & Brady, 1982; Woodall & Burgoon, 1984). This may be attributed to the fact that, despite the advancement of phonetic technology, speech rate was one of the most easily manipulable voice parameters, both subjectively and objectively.

The *fourth period* (since 2000) represents contemporary vocal personality perception research focusing on investigating gender (stereotypes) and voice qualities. While early research examined predominantly male voices (Brown et al., 1972, 1973, 1974; Moore, 1939; Scherer, 1978), recent studies employed voices of both genders, focused exclusively on female voices (e.g., Borkowska & Pawlowski, 2011; Krahé & Papakonstantinou, 2020; Levitt & Lucas, 2018; van Borsel et al., 2009), or investigated how listeners of the opposite sex perceive personality through voice (e.g., Gocsál, 2009; Jones et al., 2010). During this period, speech processing and analysis software was constantly developed to allow researchers to simplify complex speech analyses and extend their work, promoting interdisciplinary research. For example, the linguistic computer program PRAAT was first launched in 1992 but gained prominence at the beginning of 2001 with more than 5,000 registered users in 99 countries (Boersma & van

Heuven, 2001). PRAAT enables speech analysis, synthesis, and manipulation and has become the standard tool in linguistic research (Boersma & van Heuven, 2001). Due to technological advancements, voice qualities, such as creakiness and breathiness, could be manipulated, resulting in the rise of such research (e.g., Pearsell & Pape, 2023; Waaramaa et al., 2021; Yuasa, 2010). Moreover, in this period, perceptual studies were extended to domains like marketing, politics, and business. For instance, findings on the perception of personality through voice were related to voting behavior (Klofstad et al., 2015; Tigue et al., 2012), success in job interviews (Anderson et al., 2014), or brand equity (Zoghaib, 2017).

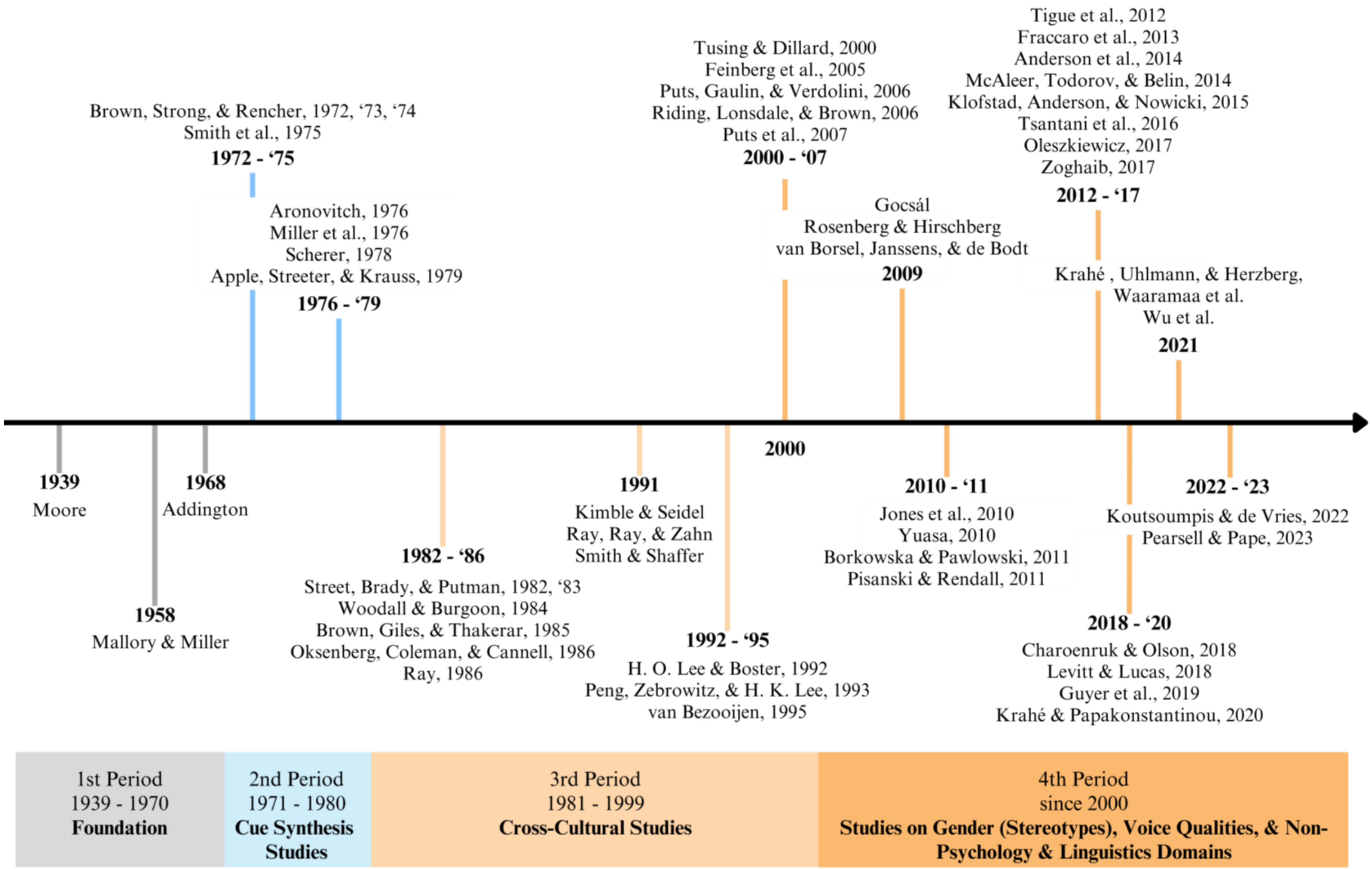


Figure 5. Evolution of Research on Vocal Personality Perception

Notes: N = 52.

5.2. Summary of Extracted Data

The 52 selected studies provide 444 single results on the relationship between voice parameters and personalities. **Table 2** illustrates the macro-level analysis of the findings categorized according to the concept matrix of extracted data.

Table 2. Macro-Level Analysis of Extracted Data

Category	Extracted Data	Macro-Level Analysis
Vocal Personality Perception	Voice Parameters	In total, 22 voice parameters were investigated, with pitch (87 single results) and speaking rate (81 single results) being the most researched voice parameters.
	Personality Traits	In total, 59 different personalities were investigated, with (social) dominance (31 single results), competence (29 single results), neuroticism, and trustworthiness (both 22 single results) being the most researched ones.
	Relationship & Analysis	In total, 426 single results showed linear relationships, and 18 showed non-linear relationships (one U-shaped relationship and 17 inverted U-shaped relationships). Further, 46 studies investigated correlations between personality traits and individual voice parameters (univariate analysis), one investigated linear combinations of voice parameters (multivariate analysis), and five investigated both.
Stimulus Material	Gender of Speaker	Within the 444 single results, 203 were concerned with the female voice, and 230 were concerned with the male voice. In the studies by Moore (1939) and Miller et al. (1976), the speaker's gender was not stated. Since it was expected to recruit men for experiments, it can be assumed that males were used in these studies (eleven single results).
	Language	Perceptual studies were found in 13 different languages: American English, British English, Canadian English, Chinese, Dutch, Finnish, French, German, Hungarian, Japanese, Korean, Polish, and Scottish English. American English was by far the most researched language, with 308 out of 444 single results (i.e., 70%) focusing on this language.
	Type of Speech Data	Most of the results were derived from studies in which the stimulus material contained read text (291 single results), e.g., an excerpt of the "Rainbow Passage" (Addington, 1968; Ray, 1986). Semi-spontaneous and spontaneous speech were used in 30 and 83 single results, respectively. Further, in seven studies, speakers recorded one or several vowels (20 single results). Koutsoumpis and Vries (2022) used read text and recorded vowels as stimulus material (12 single results). In the study by Moore (1939), information on the speech data was missing (eight single results).
Methodology	Study Design	In total, 22 perceptual studies utilized a cue synthesis study design (83 single results), 15 followed a cue manipulation study design (230 single results), and another 15 applied a correlation study design (131 single results).
	Assessment Method of Personalities	Only four studies used self-ratings to assess perceived personality traits (32 single results). Thus, more than 90% of the studies applied external judgment ratings to consider for evaluating personalities through voice (412 single results).

5.3. Voice Parameters Inducing Personality Perceptions

Within the literature, 22 voice parameters that induce personality perceptions have been identified. In the following, these voice parameters are defined and explained as their understanding is critical for interpreting the micro-level analysis of the findings (see **Table 3**).

Voice parameters can be described according to their physical acoustical features, which are objectively measurable, or their subjective perception through listeners (Hildebrand et al., 2020). For example, a voice can be perceived as being rough by listeners. However, from a physical standpoint, roughness describes irregular vocal fold vibrations accompanied by random glottal pulse fluctuations, which can be measured through acoustical measures like harmonics-to-noise ratio (HNR).

Further, voice parameters can be categorized according to their soundwave dimension: timing, frequency, amplitude, and spectral features of a voice, which will also be referred to as the voice quality dimension in the remainder of this article (Clark et al., 2007; Hildebrand et al., 2020). The *timing dimension* refers to the duration or length of a soundwave. *Frequency* refers to a soundwave's frequency, measured in Hertz (Hz). The intensity level captures the *amplitude* dimension of a soundwave, represented as the amount of power per unit area and measured in decibels (dB; Hildebrand et al., 2020). The *voice quality* dimension captures perturbations of a soundwave by analyzing its spectral features and assessing the level of periodicity. Voice qualities can be described on a continuum with breathy on one end, a modal voice with a regular and periodic vibration pattern in the middle, and pressed on the other (Pompino-Marschall, 2009). Compared to the other soundwave dimensions, the dimension of voice quality can be considered a multidimensional construct, meaning that the perception of certain voice qualities, such as breathiness or roughness, results from a combination of several voice parameters (Barsties V Latoszek, Mathmann, & Neumann, 2021).

Table 3. Definition of Voice Parameters

Soundwave Dimension	Listener's Perception	Acoustical Measure (Metric)	Definition/ Explanation
Timing	Fluency	Speaking Rate (syl/s)	Number of words or syllables spoken per unit of time, including voice breaks (Tusing & Dillard, 2000).
		Voice Breaks (n) Length of Pauses (s)	Voice breaks refer to pauses made during speech. Two types of pauses exist: filled pauses are voice breaks that are filled by a vocalization, e.g., "um" or "eh"; unfilled pauses are silent voice breaks and might be accompanied by inspiration, expiration, or swallowing (Conrad et al., 2008; Koutsoumpis & Vries, 2022; Rodero, 2012).
Frequency	Pitch	Fundamental Frequency Mean (f0 Mean; Hz)	Pitch is the perceptual representation of the fundamental frequency (f0), the number of vibrations per second that the vocal folds make to produce a vocalization (Hildebrand et al., 2020; Koutsoumpis & Vries, 2022).
	Pitch Variability	Fundamental Frequency Standard Deviation (f0 SD; st) Glide (Hz)	Pitch variability refers to the pitch rise and fall over an utterance and is measured as the standard deviation of f0 or glide. Pitch variability is used as an index of intonation, describing how monotone (narrow pitch variability) or dynamic (wide pitch variability) voice is perceived (Hodges-Simeon et al., 2010; Koutsoumpis & Vries, 2022). Glide is measured as f0-end minus f0-start (McAleer et al., 2014).
Amplitude	Loudness	Intensity Level (dB)	Loudness is the perceptual representation of a voice's amplitude/ intensity level. The perceived loudness increases (decreases) as the intensity level increases (decreases). Speakers can vary their sound energy by changing lung pressure and vocal fold opening (Hodges-Simeon et al., 2010; Tusing & Dillard, 2000).
	Loudness Variability	Intensity Standard Deviation (SD; dB) Intensity Range (dB)	Loudness variability can be used as an index for intonation and can be described through the standard deviation and range of the intensity while speaking. The intensity range is the difference between the highest and lowest intensity levels during speech. A wide (narrow) intensity range indicates a dynamic (monotone) intonation (Scherer, 1974).
Voice Quality	Breathiness	Alpha Ratio/ Spectral Slope	Breathy voices occur due to inefficient vocal fold vibration, resulting in incomplete fold closure caused by minimal laryngeal tension. As a result, the voice becomes mixed with unmodulated transglottal airflow, leading to the recognizable turbulent noise (Gobl & Chasaide, 2003; Klasmeyer & Sendlmeier, 1997). The alpha ratio or spectral slope is calculated as the energy difference between two freely selectable frequency bands (McAleer et al., 2014). A single slope value is hardly meaningful; however, this parameter indicates breathiness, as breathy voices have steeper spectral slopes (M. Gordon & Ladefoged, 2001; Kuang & Liberman, 2018; McAleer et al., 2014).
	Roughness/ Hoarseness	Harmonics-to-Noise Ratio (HNR; dB)	Rough or hoarse voices are produced through irregular vocal fold vibrations accompanied by random glottal pulse fluctuations (Barsties V Latoszek, Bodt de, et al., 2018; Dejonckere et al., 1993).

			The harmonics-to-noise ratio (HNR) represents the relationship between speech's periodic (harmonics) and aperiodic (noise) components. HNR indicates roughness or hoarseness in the voice (Anjos de Oliveira et al., 2020; McAleer et al., 2014).
	Creakiness		Creakiness describes a voice quality characterized by creaking, cracking, and popping noises. The creak is produced through the very short opening of the vocal folds and the abrupt irregular vibrations of the vocal folds. This voice quality typically occurs when speakers lower their vocal pitch by about one-half of their usual fundamental frequency (f0). A creaky voice is also known as vocal or glottal fry, laryngealization, glottalization, pulsation, pulse phonation, and trillization. (Clark et al., 2007; Klasmeyer & Sendlmeier, 1997)
	Tension		Tense or strained voices are produced through a high tension in the entire vocal tract and the larynx and are also described as "metallic" (p. 35) by Moore (1939; Laver, 1980).
	Harshness		Harsh voices can be defined as extremely tense voices. Next to the high tension in the vocal tract and larynx, harsh voices additionally show significant aperiodic vibrations of the vocal folds (Childers & Lee, 1991; Laver, 1980; Wendahl, 1963).
	Resonance/ Thinness	Formant Dispersion (Df; Hz)	Resonant or orotund voices are smooth and robust without an extensive pitch range (Zuckerman & Miyake, 1993). Thin voices can be described as the opposite of resonant voices. Formant dispersion (Df) is the ratio between successive formants (f1 to f4) and describes the characteristics of the resonance frequencies in the supralaryngeal vocal tract. Male speakers tend to have longer vocal tracts, which produce lower, more closely spaced formants and, thus, less dispersed formants, giving the voice a more resonant sound and fuller timbre. Females, on the other hand, tend to produce higher formants and formant dispersion due to their shorter and thinner vocal tracts (Levitt & Lucas, 2018; McAleer et al., 2014; Pisanski et al., 2014; Puts et al., 2011).
	Nasality		Nasality is produced by a lowered velum (soft palate), causing air to escape through the nasal cavity (Pearsell & Pape, 2023). Nasal voices are described as "nasal whine" (p. 34) or "whiny" (p. 34) by Moore (1939).
	Flatness/ Sharpness		Although Addington (1968) and Scherer (1978) have researched flat and sharp voices, they do not define these terms. Flat voices are often called monotone in everyday language, with low pitch variability. However, in his study, Addington (1968) also examined pitch variability in addition to flat voices. The results of pitch variability and flat voices are incomparable and often contradictory. It is reasonable to assume that the flat voices in Addington's (1968) study were highly unexpressive voices with low pitch variability, low loudness variability, and low resonance. While sharp voices are often described as high-pitched, Scherer (1978) also examined "high pitch". Instead, sharp voices may exhibit additional voice qualities such as high resonance. The interpretation of these voice parameters was limited due to unclear definitions.

Notes: dB = Decibel; Df = formant dispersion; f0 = fundamental frequency; Hz = Hertz; n = number; s = second; SD = standard deviation; st = semitones; syl = syllable.

6. Micro-Level Analysis

The purpose of this study is to provide marketers with guidance in determining the appropriate voice to fit a desired brand personality. Therefore, the micro-level analysis analyzes and interprets the findings on which voice parameters induce personality perceptions. In doing so, the results on the vocal perception of human personalities are screened for applicable and relevant personalities to brands, following the brand personality dimensions and traits of Aaker's (1997) BPS: sincerity, excitement, competence, sophistication, and ruggedness. Drawing on the findings, vocal profiles for these brand personality dimensions are derived to guide brand voice selection.

6.1. Scope of Analysis

For the micro-level analysis of extracted data, only results from *external judgments* and in *American English* are considered to make meaningful managerial recommendations on finding the appropriate voice for a desired brand personality.

Within the literature on vocal personality perception, American English is by far the most researched language (70% of extracted single results), which makes it possible to consolidate and compare the results. Further, even though research has shown that the dimensions of human personality are the same in different cultures, that does not automatically apply to brand personality (Aaker, 1997; Paunonen et al., 1992). Aaker's (1997) brand personality scale was developed with a U.S. sample, making the scale valid for America and the perception of brands within its culture. Following Aaker (1997), numerous studies in the field of brand personality research have shown that other scales had to be developed for individual countries or cultures and that Aaker's (1997) scale could not be generalized in Asian countries (Ferrandi et al., 2015; Sung & Tinkham, 2005).

Further, only studies with external judgments as assessment methods of personalities are considered in the following analysis, as research indicates that personality perception may differ in terms of accuracy and validity based on the assessment method (Koutsombogera et al., 2020). Only four studies (8%) used self-ratings to measure personality, making these results unsuitable for individual consideration. In addition, results on personality perception through self-ratings may not be necessary for the design of future brand voices, as it is more important how brand voices are perceived by listeners, which is reflected by external judgments.

Finally, it needs to be noted that most studies examined individual voice parameters and their (non-)linear relationship with personalities, i.e., univariate analysis of voice parameters. Only six studies examined linear combinations of several voice parameters and their joint effect on the perception of personalities, i.e., multivariate analysis. While the individual results are summarized and discussed below for the brand personality dimensions, it is essential to note that the particular results shall not be combined. For example, pitch and speaking rate individually correlate positively with a competent personality perception (Oksenberg et al., 1986). However, it does not automatically mean that the combination of pitch and speaking rate will also correlate positively with competence. The combination of a high-pitched voice and a fast speaking rate may even lead to the perception of less competence. In speech, the voice parameters covary with each other, meaning the combination of voice parameters can lead to interaction effects that cannot be assessed without further investigation (Apple et al., 1979).

6.2. Brand Personality Perception Through Voice

6.2.1. Vocal Profile of Sincerity

The personality dimension of sincerity describes personalities that are associated with warmth, acceptance, and honesty, which is why this dimension strongly relates to the human personality dimension agreeableness of the Big Five (Aaker, 1997). Traits like down-to-earth, sincere, wholesome, sentimental, and friendly represent the sincerity brand personality.

In the literature findings, none of the personality traits and their vocal perception of this brand personality dimension are found directly. Nevertheless, several personality traits were studied that can be placed in the same spectrum of the sincerity brand personality dimension: benevolence, credibility, politeness, sensitivity, and social attractiveness. Further, insights on the vocal perception of the personality dimension of agreeableness exist and are used to describe the vocal profile of sincerity due to the strong relationship between these two personalities. **Table 4** shows the vocal profiles of sincerity-related personality traits for both genders.

Based on the vocal perception of agreeableness, benevolence, credibility, politeness, sensitivity, and social attractiveness, a sincere brand personality may be perceived when **male voices** with a high pitch variability or low loudness are used (Addington, 1968; Brown et al., 1973, 1974; Ray, 1986). Various voice qualities like sharpness, thinness, tension, nasality, or inexpressiveness (i.e., flat voices) mitigate the perception of sincerity-related personality traits (Addington, 1968; Brown et al., 1985; Scherer, 1978). Contradicting results exist concerning the perception of pitch and speaking rate and should be considered when selecting the appropriate male voice for sincere brands. Even though the majority of studies identified a negative correlation between pitch and agreeableness, benevolence, and social attractiveness (Brown et al., 1974; Riding et al., 2006; Scherer, 1978), there was also an inverted U-shaped relationship found for benevolence (Riding et al., 2006). The speaking rate correlated positively with social attractiveness and credibility (S. M. Smith & Shaffer, 1991; Street & Brady, 1982; Street et al., 1983) and negatively with benevolence in different studies conducted by Brown et al. (1972, 1973, 1974). Interestingly, S. M. Smith and Shaffer (1991) found increased credibility ratings induced through fast speaking rates for counterattitudinal messages and an inverted U-shaped relationship for proattitudinal messages.[3] Apple et al. (1979) confirmed an inverted U-shaped relationship between speaking rate and neutral speech content.

Brands with the personality sincerity, which are represented by **female voices,** should use a voice that has a high pitch variability or is breathy, tense, or less resonant (positive correlation with thin voice and negative correlation with flat voice; Addington, 1968; Ray et al., 1991). Moreover, a medium speaking rate is favorable, and voice qualities like nasality or roughness should be avoided as they decrease perceptions of politeness (Addington, 1968; H. O. Lee & Boster, 1992; Ray et al., 1991).

Differences between the genders exist on the one hand because more research has been done on the male voice, and personality traits such as agreeableness, benevolence, and credibility have been studied, which are lacking in female voices. On the other hand, there are differences concerning the perception of speaking rate for social attractiveness and a tense voice for sensitivity. Men with faster speaking rates are perceived as socially attractive (H. O. Lee & Boster, 1992; Street & Brady, 1982; Street et al., 1983); for women, speaking rate indicates an inverted U-shaped relationship with social attractiveness assessments, meaning a medium speaking rate is best (Ray et al., 1991). For the perception of sensitivity, the study of Addington (1968) revealed contradicting results for tense voices – a tense male voice decreased sensitivity assessments, whereas a tense female voice increased them.

In conclusion, inferences about how a male and female brand voice with a sincere personality could sound are made by combining and interpreting findings on the vocal perception of personalities like agreeableness, benevolence, credibility, politeness, sensitivity, and social attractiveness. Different results are found for pitch and speaking rate depending on the personality traits of male voices. Furthermore, S. M. Smith and Shaffer's (1991) study showed that the

---

[3] The counterattitudinal message contained six arguments favoring a drinking age in favor of a drinking age of 21. The proattitudinal message contained six arguments against a drinking age of 21.

message's content can influence the vocal perception of sincerity-related traits. Accordingly, credibility perception increased when counterattitudinal messages were spoken at a fast speaking rate and proattitudinal and neutral messages were spoken at a medium speaking rate (S. M. Smith & Shaffer, 1991). In the case of female voices, no contradicting results on vocal perception of politeness, sensitivity, and social attractiveness exist. Nonetheless, single results are derived from only three studies, most of which came from the cue manipulation study by Addington (1968). Ensuring that other voice parameters were not changed simultaneously in cue manipulation studies is impossible. In addition, Addington (1968) also missed defining voice parameters like flat or thin voice, making it more challenging to rely on the results.

Finally, it must be noted that the analysis of the identified vocal perception of sincerity-related personality traits presented here only indicates how a sincere brand voice can sound. Consequently, more research must be conducted on the vocal perception of sincerity and brand personality traits like down-to-earth, honest, wholesome, and cheerful to provide marketers with more meaningful guidance in sincere brand voice selection.

Table 4. Vocal Profiles of Brand Personality Traits Related to Sincerity

	Male Voice			Female Voice		
	⊕	∩	⊖	⊕	∩	⊖
Agreeableness			Pitch; Sharpness; Thinness			
Benevolence*	Pitch variability	Pitch; Speaking rate	Loudness; Pitch; Speaking rate			
Credibility	Speaking rate	Speaking rate				
Politeness	Pitch variability		Nasality; Tension	Breathiness		Nasality; Roughness
Sensitivity	Pitch variability		Flatness; Tension	Tension; Thinness		Flatness
Social Attractiveness	Speaking rate		Pitch	Pitch variability	Speaking rate	

Notes: Plus = positively correlated; minus = negatively correlated; inverted U = inverted U-shaped relationship; *Studies on linear combinations of voice parameters showed that benevolence is perceived in male voices when they have a slow speaking rate together with a high pitch variability (Ray, 1986), low loudness together with high pitch variability (Ray, 1986), and low-pitch variability together with medium-pitch (Riding et al., 2006).

6.2.2. Vocal Profile of Excitement

The excitement personality dimension describes personalities associated with activity, energy, and sociability. It captures the elements of the human personality dimension extraversion of the Big Five (Aaker, 1997; Aaker et al., 2001). Brands with an exciting personality can be described as daring, cool, young, imaginative, independent, and up-to-date.

Only the vocal perception of the personality trait "young" was found in the literature, which is included in the excitement brand personality dimension. Additionally, several human personality traits are studied that could be placed in the same spectrum of the excitement brand personality dimension: activity, boldness, dynamic, enthusiasm, talkativeness, and youthfulness. Further, there are findings on the vocal perception of extraversion, a personality dimension from the Big Five, which is used to describe the vocal profile of excitement due to the strong relationship between these two personality constructs. **Table 5** shows the vocal profiles of excitement-related personality traits for both genders.

Based on the vocal perception of activity, boldness, enthusiasm, extraversion, talkativeness, and young, the brand personality excitement might be perceived when **male voices** with a high pitch, high loudness, or a dynamic intonation (indicated by a positive correlation with pitch variability and loudness variability and negative correlation with flat voice) are used (Addington, 1968; Aronovitch, 1976; Scherer, 1978). Additionally, the voice should be resonant or breathy but less rough or tense to be assessed with excitement (Addington, 1968; Scherer, 1978). A fast speaking rate leads to the perception of activity, enthusiasm, extraversion, and talkativeness but decreases the perception of boldness (Aronovitch, 1976). Nasality seems to decrease the perception of activity and enthusiasm (Addington, 1968), but increases the perception of extraversion (Scherer, 1978). Thus, marketers should consider a tradeoff in speaking rate and nasality adjustments when selecting a male voice for an exciting brand. The specifications of these voice parameters must be adjusted depending on which personality trait is to be communicated.

Differences between male and female voice profiles regarding the excitement brand personality are minimal. Like male voices, **female voices** should be loud or have a dynamic intonation (indicated by a positive correlation with pitch variability and a negative correlation with flat voice; Addington, 1968; Aronovitch, 1976; Ray et al., 1991). Additionally, a fast speaking rate, breathiness, tension, or resonance increases the perception of excitement (Addington, 1968; Aronovitch, 1976; Ray et al., 1991). Voices with roughness, nasality, or creakiness should be avoided as they decrease assessments of excitement-related personality traits like activity, enthusiasm, talkativeness, and youthfulness (Addington, 1968; Levitt & Lucas, 2018). The positive correlation between thin voices (opposite of resonant voices) and the personality trait “young” contradicts the positive correlations between resonance and formant dispersion with personality traits like enthusiasm, talkativeness, and youthfulness (Addington, 1968). Thus, marketers who want to select a voice for an exciting brand personality that should sound young might choose less resonant voices, as resonance could indicate maturity.

In summary, like brand personality sincerity, there is limited research on the brand personality traits of excitement, such as daring, spirited, imaginative, and up-to-date. Direct results are only available for the excitement personality trait “young” derived from only one study on voice quality parameters, that of Addington (1968). As mentioned earlier, in the cue manipulation study by Addington (1968), it was impossible to ensure that other voice parameters were kept the same simultaneously. It is, therefore, advisable to look at the results for other excitement-related personality traits like activity, boldness, dynamic, enthusiasm, talkativeness, youthfulness, and extraversion.

Both gender voice profiles indicate that a brand with an exciting brand personality should be high in loudness or have a dynamic intonation. The genders differed in that male voices with a high pitch are perceived as bold and extraverted, whereas no results on female voice pitch and excitement-related personality traits exist. Conversely, results on the perception of female speaking rate and excitement-related personality traits are more consistent (correlations are all positive), whereby they are slightly inconsistent for male voices (negative correlation only with boldness). In conclusion, marketers are advised to consider the described voice profiles of excitement as an indicator. However, as more research is needed on the vocal perception of the brand personality dimension excitement, marketers should pay closer attention to individual excitement-related personality traits when selecting a suitable exciting brand voice.

Table 5. Vocal Profiles of Brand Personality Traits Related to Excitement

	Male Voice		Female Voice	
	+	−	+	−
Activity/ Energy	Speaking rate; Pitch variability; Resonance	Flatness; Nasality	Speaking rate; Loudness; Tension	Flatness; Roughness
Boldness	Loudness; Pitch	Speaking rate	Loudness; Speaking rate	
Dynamic			Pitch variability; Speaking rate	
Enthusiasm	Speaking rate; Pitch variability; Resonance	Flatness; Nasality	Speaking rate; Pitch variability; Resonance; Tension; Breathiness	Flatness; Nasality; Roughness
Extraversion	Loudness; Loudness variability; Nasality; Pitch; Speaking rate		Loudness; Speaking rate	
Talkativeness	Pitch variability; Speaking rate	Flatness	Resonance; Pitch variability; Speaking rate	Roughness
Young	Breathiness	Roughness; Tension	Breathiness; Tension; Thinness	
Youthfulness			Formant dispersion	Creakiness

Notes: Plus sign = positively correlated; minus sign = negatively correlated.

6.2.3. Vocal Profile of Competence

The competence personality dimension describes personalities associated with responsibility, dependability, and achievement, attributes that are also present in the human personality dimension conscientiousness of the Big Five (Aaker, 1997; Aaker et al., 2001). Competent brands are described through intelligent, confident, reliable, successful, and corporate attributes. Unlike the previous personalities, 12 studies provided direct results on perceiving competence through voice parameters. **Table 6** shows both genders' vocal profiles of competence and related personality traits.

A brand with a competent personality is best perceived through a **male voice** with high loudness, high pitch variability, low pitch, or a fast speaking rate (Brown et al., 1972, 1973, 1974; Klofstad et al., 2015; H. O. Lee & Boster, 1992; Peng et al., 1993; Ray, 1986; B. L. Smith et al., 1975; Street & Brady, 1982; Street et al., 1983). Surprisingly, nine studies differing in study design and type of speech data showed positive correlations between speaking rate and competence assessments. In addition, Ray (1986) investigated pitch and speaking rate as a linear combination for the perception of competence and found that a high pitch variability and a fast speaking rate increased competence perception. In contrast, a low pitch variability and a slow speaking rate led to the opposite.

There is a substantial similarity between the female and male voice profiles. **Female voices** should also be high in loudness and pitch variability or have a fast speaking rate to be perceived as competent (Oksenberg et al., 1986). A creaky voice should be avoided, as creakiness mitigates the perception of competence (Anderson et al., 2014). Contradicting results exist concerning the perception of competence through pitch in female voices. The cue synthesis study of Klofstad et al. (2015) indicated a negative correlation between pitch and competence, meaning female low-pitched voices are perceived as more competent than high-pitched voices, which would follow the male voice profile. In contrast, Oksenberg et al.'s (1986) correlation study showed that female high-pitched voices are perceived as more competent. This difference reveals that the perception of female voices (in contrast to male voices) is more complex regarding competence. On the other hand, the difference can also be due to the diverse vocal stimuli, study designs, or speech content.[4] Either way, the conflicting results indicate that more research is needed on female pitch and perception of competence.

[4] Participants of Klofstad et al.'s (2015) study listened to five speakers saying the sentence, "I urge you to vote for me this November." (p. 4). Oksenberg et al.'s (1986) study participants listened to ten speakers in an introductory paragraph of a telephone interview.

As with the other personality dimensions, personalities linked to Aaker's (1997) competence brand personality dimensions like charisma, (self-)confidence, intelligence, knowledgeability, reliability, trustworthiness, and willingness to cooperate are also considered. The following discusses only the differences between competence and competence-related personality traits (see **Table 6**). This discussion provides a deeper understanding of how nuances in personalities that are closely related yet different are crucial in vocal personality perception and highlights potential areas for future research.

Similar to the perception of competence, loudness and speaking rate show several positive correlations with most competence-related personality traits, e.g., charisma,[5] confidence, and intelligence, in **male voices** (Aronovitch, 1976; Charoenruk & Olson, 2018; H. O. Lee & Boster, 1992; Miller et al., 1976; Rosenberg & Hirschberg, 2009). Even though Brown et al. (1974) and Klofstad et al. (2015) revealed that competence is perceived through low-pitched male voices, charisma and self-confidence are found to be perceived through high-pitched male voices (Aronovitch, 1976; Rosenberg & Hirschberg, 2009). Further, Charoenruk and Olson (2018) indicated that confidence, reliability, and trustworthiness ratings are highest for a medium pitch, i.e., inverted U-shaped relationship, which means neither particularly low nor high male voices are perceived as "competent". Additionally, the same study contradicted the positive correlation of pitch variability with competence found by Brown et al. (1972, 1974) and Ray (1968) by finding negative correlations of pitch variability with confidence, reliability, and trustworthiness ratings. Since Charoenruk and Olson's (2018) study is the only one that contradicted the positive correlations of pitch variability and competence, marketers should not rely heavily on its results when selecting a male voice for a competent brand personality. Further significant voice parameters in the perception of competence-related personality traits might be

[5] Charisma is a multidimensional construct that combines several personalities. Because charisma clusters personality traits and attributes like leadership, persuasion, confidence, and authority, it is used as a personality related to the competence brand dimension (Rosenberg & Hirschberg, 2009; Signorello et al., 2012).

helpful in voice design for a competent male brand. For example, voice breaks like pauses in speaking diminish the perception of confidence and reliability, and voice qualities like nasality or tension reduce the perception of intelligence and willingness to cooperate (Addington, 1968; Charoenruk & Olson, 2018).

Similar to the male voice, also loudness and speaking rate in **female voices** show several positive correlations with most of the competence-related personality traits, e.g., self-confidence, reliability, and willingness to cooperate (Addington, 1968; Aronovitch, 1976; Charoenruk & Olson, 2018). As mentioned earlier, contradictory results exist concerning the perception of competence through pitch in female voices since competence assessments showed a positive and negative correlation with pitch in two studies (Klofstad et al., 2015; Oksenberg et al., 1986). Charoenruk and Olson (2018) indicated inverted U-shaped relationships between pitch and confidence, reliability, and trustworthiness, like for the male voices. While one cannot wholly rely on the results of a single study, the identified inverted U-shaped relationships between pitch and competence-related personalities again point to the need for more research into the perception of female pitches with competence and related personalities.

In summary, when selecting a suitable female voice for a competent brand personality, marketers should rely primarily on the results of the competence perception study of Oksenberg et al. (1986). Additional voice parameters are significant for perceiving competence-related personality traits in female voices. For example, voice breaks decrease the perception of confidence, reliability, and trustworthiness (Charoenruk & Olson, 2018). Furthermore, creakiness negatively impacts the perception of competence and related personalities, such as confidence, intelligence, and trustworthiness (Anderson et al., 2014; Yuasa, 2010). Especially the perception of intelligence, which is included as a personality trait of Aaker's (1997) competence dimension, seems to be negatively influenced by several voice qualities like nasality, roughness, tension, and thinness (Addington, 1968).

Table 6. Vocal Profiles of Brand Personality Competence and Related Personality Traits

	Male Voice			**Female Voice**		
	+	∩	−	+	∩	−
Competence*	Loudness; Pitch variability; Speaking rate		Pitch	Loudness: Pitch; Pitch variability; Speaking rate		Creakiness; Pitch
Charisma	Loudness; Pitch					
Confidence	Speaking rate	Pitch	Pitch variability; Voice breaks	Speaking rate	Pitch; Pitch variability	Creakiness; Voice breaks
Intelligence	Speaking rate		Nasality; Pitch; Tension			Creakiness; Nasality; Roughness; Tension; Thinness
Knowledgeability	Speaking rate					
Reliability	Speaking rate	Pitch	Pitch variability; Voice breaks	Speaking rate	Pitch	Pitch variability; Voice breaks
Self-confidence	Loudness variability; Pitch; Speaking rate			Loudness; Speaking rate		
Trustworthiness	Speaking rate	Pitch	Pitch; Pitch variability	Speaking rate	Pitch; Pitch variability	Creakiness; Voice breaks
Willingness to Cooperate	Pitch variability		Tension	Speaking rate		

Notes: Plus sign = positively correlated; minus sign = negatively correlated; inverted U = inverted U-shaped relationship; *Studies on linear combinations of voice parameters showed that competence is perceived in male voices, which are high in pitch variability and have a fast speaking rate (Ray, 1986).

6.2.4. Vocal Profile of Sophistication

The following two brand personalities, sophistication and ruggedness, differ from any of the Big Five, indicating that brand personality perceptions differ from human perceptions. Aaker (1997) explained that sophistication and ruggedness represent personalities humans de-sire yet do not possess. Brands with these personalities typically exaggerate characteristics such as “sexiness” and “glamor” for prototypical sophisticated brands or “strength” and “masculinity” for rugged brands. Such extremes tend to be the exception with human personalities.

Sophistication as brand personality is attributed to traits like charming, smooth, feminine, upper class, glamorous, and good-looking. Within the literature findings, only one study provided direct results on the perception of sophistication through voice parameters (Addington, 1968). In addition, the same study gave insights into the perception of femininity, a trait of the sophisticated brand personality. **Table 7** shows the vocal profiles of sophistication and the related personality traits of feminity for both genders.

According to the cue manipulation study of Addington (1968), who mainly investigated the perception of voice qualities on personality traits, sophistication is perceived through resonant and rough **male voices**. Further, nasality is found to mitigate sophistication assessments in males. For **female voices**, results show that various voice qualities like breathiness, nasality, roughness, tension, and thinness decrease the perception of sophistication.

Concerning the perception of feminity, results for **female voices** contradict those of sophistication perceptions since breathiness, tension, and thinness increase feminity assessments. Also, rough and flat voices, i.e., low intonation and less resonance, lead to negative perceptions of feminity. Feminity assessments increase for **male voices** with a high pitch variability. Results on the influence of voice qualities on the perception of feminity are insignificant for male voices.

In conclusion, even though direct results on the perception of sophistication exist, they derive only from one study, the cue manipulation study conducted by Addington (1968). As already mentioned, given the type of study design, it is possible that the speakers changed other voice qualities at the same time, as they were “instructed to simulate seven voice qualities […] and three speaking rates” (Addington, 1968, p. 493). Accordingly, the results can be used as an indicator for selecting a voice representing a sophisticated brand without relying on them entirely. The limited number of studies and the contradictory results in the perception of sophistication and feminity through vocal qualities in the female voice show that more research is needed for this brand personality dimension. Further, insights on sophistication perception through voice parameters of the other voice dimensions, timing, frequency, and amplitude, are missing. This lack of research into the perception of sophistication and related personalities by voice confirms that the vocal perception of human personalities, particularly the Big Five, has been predominantly studied.

Table 7. Vocal Profiles of Brand Personality Sophistication and Related Personality Trait

	Male Voice		Female Voice	
	+	−	+	−
Sophistication	Resonance; Roughness	Nasality		Breathiness; Nasality; Roughness; Tension; Thinness
Feminity	Pitch variability		Breathiness; Tension; Thinness	Flatness; Roughness

Notes: Plus sign = positively correlated; minus sign = negatively correlated.

6.2.5. Vocal Profile of Ruggedness

The fifth brand personality dimension of Aaker's (1997) BPS is ruggedness. Brands with this personality are attributed characteristics such as outdoorsy, masculine, western, tough, and rugged. As already mentioned, ruggedness bears no resemblance to any of the human personality dimensions of the Big Five (Aaker, 1997; Aaker et al., 2001). Within the literature, only findings on the vocal perception of masculinity, included as a trait in the dimension of ruggedness, are found for male voices. No other personality trait appears to relate to the brand personality of ruggedness. **Table 8**, therefore, shows the voice profiles of masculinity for male voices as the only indicator of a rugged brand personality.

Insights on the perception of masculinity through male voice are derived from two perceptual studies. According to Feinberg et al. (2005), masculinity is perceived when a **male voice** is low-pitched or shows less dispersed formants, which leads to a more resonant sound and fuller timbre. In addition, Feinberg et al. (2005) examined pitch and formant dispersion as a linear combination for the perception of masculinity. They found that a low pitch and a low formant dispersion increase the perception of masculinity individually and together. According to Addington (1968), flat voices increase masculinity assessments. Suppose it can be assumed that flatness describes voices with low pitch variability, low loudness variability, and low resonance. In that case, the studies of Feinberg et al. (2005) and Addington (1968) contradict each other regarding resonance. However, as Addington (1968) did not define a flat voice, the perception of masculinity through this voice quality must be interpreted with care.

Since few results exist on ruggedness or ruggedness-related personalities, it is impossible to reasonably interpret the findings and make recommendations for marketers on selecting a rugged brand voice. Especially for female voices, insights on the vocal perception of ruggedness-related personality traits are missing. In conclusion, more research on this brand personality dimension is needed to give brands with rugged personalities the appropriate voice.

Table 8. Male Voice Profile of Brand Personality Trait Masculinity

	Male Voice	
	+	−
Masculinity*	Flatness	Formant dispersion; Pitch

Notes: Plus sign = positively correlated; minus sign = negatively correlated; *Studies on linear combinations of voice parameters showed that masculinity is perceived in male voices, which are low in pitch and have a low formant dispersion (Feinberg et al., 2005).

7. Future Research Agenda

The identified research gaps within the literature on vocal perception of human and brand personality are used to propose a future research agenda. The lack of insights on the vocal perception of brand personalities represents the most significant recommendation for the future: *research is generally needed into the perception of brand personalities through voices*. Even though brand and human personalities relate in some respect to each other, brands can embody and exaggerate characteristics, which are an exception in human personalities. Thus, marketing research has investigated specific personality dimensions and traits that apply to brands and developed BPS for different products and services, e.g., consumer goods, tourist destinations, or stores, and different cultures and languages, e.g., USA, Spain, or South Korea (Aaker et al., 2001; d'Astous & Lévesque, 2003; Grohmann, 2009; H. Lee & Cho, 2017; Usakli & Baloglu, 2011). In this regard, it would also be interesting to find out whether and which brand personalities developed by Aaker (1997) or others, like Geuens et al. (2009), Ferrandi et al. (2015), or Chu and Sung (2011), can be represented vocally.

As already mentioned in the analysis of the findings at the micro level, deriving findings on human personality traits for brand personalities is only possible to a limited extent. The results on the vocal perception of personality traits strongly related to the respective brand personality dimension are used to guide brand owners in designing or selecting a suitable spokesperson for their brand-consumer communication through voice-based devices. Direct conclusions are drawn for the brand personality competence and sophistication, as empirical findings are available for these personalities based on vocal perception in the human voice. For the brand personality dimensions of sincerity and excitement, conclusions are drawn primarily from the voice perception of agreeableness and extraversion, the Big Five dimensions of human personality, which correlate strongly with these brand personalities. Recommendations on how a brand with a rugged personality should sound cannot be given, as no results on personality traits of this dimension are found, except for the perception of masculinity in male voices.

In addition to this general lack of research in brand personality perception through voice, further recommendations are derived based on the findings, which refer to both vocal perceptions of brands and human personalities (see **Figure 6**).

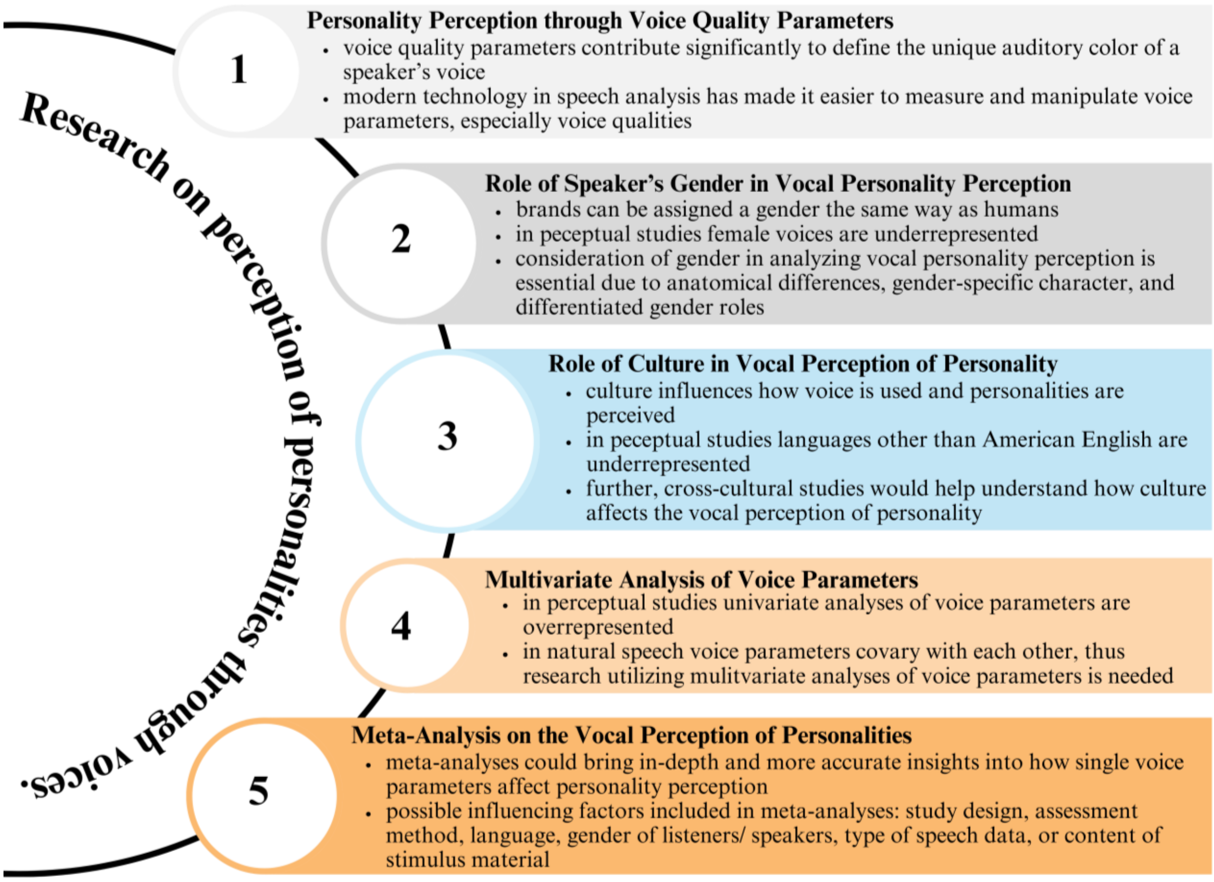


Figure 6. Avenues for Future Research on Vocal Personality Perception

7.1. Personality Perception through Voice Quality Parameters

Within the reviewed perceptual studies, 22 voice parameters are investigated, with pitch (87 single results) being the most researched. Pitch is regarded as one of the most critical vocal cues used to identify different characteristics of people, including age, gender, and mood, and it plays an essential role in people's interactions and decisions (D. Y. Huang et al., 2009; Tsantani et al., 2016).

Of the 444 single results, 181 deal with the perception of voice quality parameters like breathiness, creakiness, or nasality (external judgments counted only). Although these findings account for 41% of the total findings, 104 out of the 181 results (57%) on the perception of voice quality parameters are derived from a single study, namely that of Addington (1968). As has been emphasized several times within this study, caution should be exercised when considering Addington's (1968) cue manipulation study, particularly concerning analyzing the brand personality sincerity and sophistication. Most results on the perception of sincerity-related personality traits, such as politeness, sensitivity, and social attractiveness, for female voices and on the perception of sophistication for both genders derive from Addington's (1968) study. Ensuring that other voice parameters are not changed simultaneously in cue manipulation studies is impossible. Furthermore, the challenge of relying on the results is exacerbated by Addington's (1968) undefined voice parameters, such as a flat or thin voice.

Thus, research on the brand personality perception of voice quality parameters like breathiness, creakiness, roughness, or hoarseness is needed, as these parameters contribute significantly to defining the unique auditory color of a speaker's voice (Pearsell & Pape, 2023). Modern technology in speech analysis has made it easier to measure and manipulate voice qualities, eliminating the need for subjective evaluations that were once necessary (Hildebrand et al., 2020; Klasmeyer & Sendlmeier, 1997). Still, the measurement of voice qualities remains challenging, as voice quality parameters are understood as the combination of several parameters that additionally correlate with each other (Barsties V Latoszek, Mathmann, & Neumann, 2021; Klasmeyer & Sendlmeier, 1997). For example, perturbations in pitch (i.e., jitter) typically appear with perturbations in loudness (i.e., shimmer); depending on which parameter is more prominent, jitter and shimmer are perceived as breathiness, roughness, or hoarseness in the voice. Despite these difficulties in measuring voice quality, several acoustic indicators and indices have been developed in recent years to assess voice qualities, e.g., the Acoustic Breathiness Index (ABI) and Acoustic Voice Quality Index (AVQI) can be used as a measure for evaluating breathiness and hoarseness, respectively (Barsties V Latoszek, Kim, et al., 2021; Barsties V Latoszek et al., 2020). These findings from phonetic research should be implemented and utilized in vocal brand and human personality perception.

7.2. Role of Speaker's Gender in Vocal Personality Perception

In numerous reviewed studies, the analyses of the perception of personality through voice parameters are conducted separately for female and male voices, or gender is included in the analyses as an influencing factor (McAleer et al., 2014; Waaramaa et al., 2021). The consideration of gender in analyzing vocal personality perception is essential for three reasons. First, due to anatomical differences in the vocal tracts, males and females produce varying voices and possess different audible characteristics. Second, gender-specific abilities and character traits are still attributed to men and women due to socially or culturally constructed norms and values, influencing how they speak and, thus, how their personalities are perceived (Addington, 1968). For instance, prepubertal children's voices may convey their gender, although there are no anatomical differences in the vocal tract between genders at this age. As these differences only emerge during puberty, children's gender perception is more likely linked to cultural or societal influences on speech development (Belin et al., 2004; Fitch & Giedd, 1999). Third, as gender roles become less differentiated, and females especially become more emancipated in many cultures, changes in how females sound and speak are expected and were already observed. For example, in recent years, young adult females in the U.S. have become accustomed to speaking with a creaky voice, especially in traditionally male-dominated industries such as finance (Anderson et al., 2014; Yuasa, 2010). This development is presumably due to women wanting to adapt to the male way of speaking and using a vocal fry to be perceived as competent, serious, or intelligent (Anderson et al., 2014; Borkowska & Pawlowski, 2011; Puts et al., 2007).

Grohmann's (2009) study on the gender dimensions of brand personalities demonstrates that brands can be assigned a gender the same way as human beings. The study developed a two-dimensional scale to measure the masculinity and feminity of a brand, which can be used in conjunction with other BPSs, such as Aaker's (1997) five dimensions of brand personality. According to Grohman (2009), a fit between the parent brand personality and the brand gender should be sought to maximize the impact of the brand-consumer relationship. Thus, insights on gender-specific vocal personality perception can be transferred to brands.

Despite the differences in gender perceptions of personality, more individual results exist for male voices in perceptual research (see **Table 9**). Of the 52 reviewed studies, 21 used only male voices as their stimulus material, whereas only seven focused solely on feminine voices; the rest used voices from both genders. On the one hand, the surplus of research on male voices may have been because, in the past, scientific research on men was simply the norm. On the other hand, some studies assumed that results on male voices were transferrable to female voices, like the study of Street and Brady (1982). In three of the four vocal dimensions, there are almost twice as many results for male as for female voices. Only for the voice quality dimension do more results for female voices exist. As future recommendations, studies on vocal personality perception should employ a more balanced approach by including voices of both genders, rather than focusing exclusively on either male or female voices. In particular, research on the perception of female voices is currently underrepresented, a gap that must be addressed.

Table 9. Overview of Gender-Specific Studies and Single Results

	Male Voice	Female Voice
No. of Studies	45	31
No. of Single Results on Voice Parameters	241	203
Frequency Voice Parameters	77	51
Intensity Voice Parameters	19	15
Timing Voice Parameters	61	30
Voice Quality Voice Parameters	90	109

Notes: Only studies that used external judgments as an assessment method are included in the statistics. Even though the speaker's gender in the studies of Moore (1939) & Miller et al. (1976) is not stated, it is assumed that male voices are used in these studies, which is why their results are counted as male voices here. As in several studies the combinations of voice parameters are investigated, the aggregated number of single results per voice parameter exceeds the total number of single results on voice parameters (241 for male voices and 203 for female voices).

7.3. Role of Culture in Vocal Perception of Personality

Since 70% of the extracted data is on vocal personality perception in American English speech contexts, the micro-level analysis focuses on interpreting the findings for this language. A reason for the underrepresentation of other languages could be the general predominance of the (American) English language and the relevance of WEIRD countries (western, educated, industrialized, rich, and democratic), like Europe and the U.S., in scientific publications (Atari & Henrich, 2023; Blasi et al., 2022; Bryant, 2022). Country- or language-specific studies on vocal personality perception might be missing in fulfilling the international scope of scientific journals, and they are not published in peer-reviewed journals but in conference proceedings. For example, during the literature search, six conference papers were identified that investigated vocal personality perceptions in German, Chinese, Dutch, British, and Canadian English (Chen et al., 2001; Michalsky et al., 2020; Poon et al., 2018; Weiss & Burkhardt, 2010, 2012). None of the authors of these six conference papers published their work in a journal that met the scope of this literature review, meaning their results were not included in the extracted data.

Even though the anatomy of the vocal tract constraints the sound produced, culture influences how people use their voice, affecting how personalities are perceived through the voice (Krauss et al., 2002; Waaramaa et al., 2021). For example, a speaker's pitch range is constrained by the vocal tract length, but where females and men place their voices within this range can be determined through cultural norms. In Japan, for instance, women are expected to speak more politely than men. A high pitch expresses this politeness in the voice. Therefore, cultural norms have conditioned female Japanese to talk in a higher pitch despite having a natural medium pitch (Krauss et al., 2002). Another example of cultural conditioning in voice production and perception refers to the average speaking rate, which differs depending on the language. For example, the average speaking rate in Spanish is approximately 7.82 syllables per second, while the average speaking rate in German is slower at approximately 5.97 syllables per second (Rodero, 2012). Over time, Spaniards residing in Germany would adjust their average speaking rate to the German pace to adapt culturally.

It is, therefore, essential to conduct further research on the perception of vocal personalities in languages other than American English. In this context, conducting a comparative analysis of the vocal perception of personality in different cultures and languages would be beneficial. Such cross-cultural studies would assess the generalizability of empirical findings and contribute to a more comprehensive understanding of the cultural influence on vocal brand personality perception. The findings would assist in understanding which voice parameters are most decisive in perceiving specific personality traits in different cultures. Using culture- or language-specific BPSs is crucial in considering the vocal perception of brand personalities. It has been proven that additional or other culture-specific dimensions exist next to brand personality dimensions that apply to several countries. For example, Aaker's (1997) sincerity, excitement, and sophistication brand personality dimensions shared similar meanings in Japan, Spain, and the United States. The brand personality ruggedness could not be confirmed in Japan and Spain, two harmony-oriented cultures. Instead, the brand personality dimension peacefulness, with traits like naïve and mild-mannered, could be identified in Japan and Spain, and the dimension passion, with traits like mystical and fervent, applied in Spain. (Aaker et al., 2001).

In studying vocal brand personality perception in different languages, it could be interesting to find out not only how natives perceive voices of their language but also how listeners of another language perceive personalities. For example, Waaramaa et al. (2021) investigated how Arabic-speaking listeners perceive Finnish voices and compared their personality perception with that of Finnish-speaking listeners. Their results show that generally, both listener groups perceive voice qualities as similar, and similar stereotypical tendencies exist. Nonetheless, their study also reveals that these two languages' perception of breathy and tense voices is the opposite (Waaramaa et al., 2021). Further, psychologist K. R. Scherer (1972, 1974) conducted cross-cultural studies of vocal personality perception, which showed that American and German listeners are highly agreeable in their perceptions of the incompetence, weakness, and warmth dimensions of American and German speakers.

7.4. Multivariate Analysis of Voice Parameters

Most of the studies within this literature review identify a linear relationship between single voice parameters and personality perception, i.e., univariate analysis of voice parameters. In summary, 46 perceptual studies investigate correlations between personality traits and individual voice parameters, one investigates linear combinations of voice parameters, i.e., multivariate analysis, and five investigate both.

Since voice parameters are known to correlate with each other and their combined effect on personality perception would be interesting to study, it is surprising that most research focused on univariate analyses of voice parameters. In early phonetic studies, such as the study of Addington (1968), investigating linear combinations of voice parameters on personality perception was already addressed as further research implication (Ray, 1986). As mentioned earlier, especially the voice dimension, voice quality can be understood as the combination of several parameters that correlate with each other (Barsties V Latoszek, Kim, et al., 2021; Klasmeyer & Sendlmeier, 1997). In their study on the effect of voice quality on emotion, mood, and attitude, Gobl and Chasaide (2003) addressed this issue and assumed that “methodological constraints” (p. 207) have so far been why multivariate analyses of voice parameters have been lacking or have failed. Addington (1968) described the investigation of several voice parameters as one of the most important but also “most difficult problems” (p. 503) for researchers. With the advancement of speech processing and analysis software and the recent developments in machine learning, this problem should be addressed in future research.

7.5. Meta-Analysis on the Vocal Perception of Personalities

This study was the first to comprehensively aggregate empirical findings on personality perception through voice and derive insights into vocal human personality perception on brands. This summary was done on a qualitative basis through a systematic literature review. As a further step, a meta-analysis on the vocal perception of personalities can bring in-depth and more accurate insights into how single voice parameters affect personality perception. Meta-analyses can focus on the perception of individual voice parameters on personalities. For example, C. J. Carpenter (2012) conducted a meta-analysis examining the influence of the speaker’s (dis)fluencies on perceptions of competence and persuasion. A pitch and speaking rate perception meta-analysis would be possible since these two voice parameters were the most researched within the literature. Another approach would be to focus on individual personality traits or dimensions, such as the most researched personalities (social) dominance, competence, neuroticism, or trustworthiness, and examine which (combinations of) voice parameters play a decisive role in their perception.

In the meta-analysis, further influencing factors extracted within this literature review can be included to determine their impact on the relationship between voice parameters and personality ratings. These are, e.g., the study design, assessment method, language, gender of listeners and speakers, type of speech, or the content of stimulus material. As described above, the speaker's gender and language/culture are found to influence personality perceptions through voice (Addington, 1968; Krauss et al., 2002; Waaramaa et al., 2021).

Further, in the perception of sincerity-related personality traits, it is highlighted that the content of the speech might moderate the perception since the study of S. M. Smith and Shaffer (1991) shows that credibility perception increases when counterattitudinal messages are spoken at an increased speaking rate and proattitudinal and neutral messages are spoken at a medium speaking rate. Another study by O'Connor and Barclay (2018) examined how socially relevant cues in male speech influence their trustworthiness and attractiveness ratings by females. Their results show that women perceive male low-pitched voices as more trustworthy and attractive when they use prosocial words instead of antisocial words.[6]

Finally, also the gender of listeners might play a role in the vocal perception of personalities as previous research demonstrates that men and women perceive the pitch of the opposite sex differently from that of the same sex (Borkowska & Pawlowski, 2011; Collins & Missing, 2003; Feinberg et al., 2008; Fraccaro et al., 2013). On the other hand, several studies tested for a moderating effect of the listener's gender, which was found to be insignificant (Kimble & Seidel, 1991; McAleer & Belin, 2019). Thus, it should be clarified whether and when the listener's gender influences the perception of personalities.

8. Summary and Conclusion

Disembodied voices have become integral to daily lives as conversational agents, voice assistants, and other voice-based AI devices have increased exponentially (voicebot.ai & Business Wire, 2020). Businesses and brands use these voice technologies for personalized and convenient customer experiences. Moreover, contemporary voice technologies hold significant marketing potential, especially regarding the social and natural way consumers interact with

[6] prosocial words: caring, fair, honest, helpful; antisocial words: cheater, fraud, liar, corrupt

brands (Jurafsky & Martin, 2020; V. Kumar et al., 2016). However, it also challenges marketers concerning brand voice selection, as visual content on voice-enabled devices and platforms is usually missing (Packard & Berger, 2024). To improve the customer experience and stand out from the competition, selecting an appropriate and distinctive voice that reflects the brand's personality is crucial. Despite the wealth of empirical evidence on the perception of personality through voice, a comprehensive summary of the findings is lacking. The question for marketers of which voice characteristics are essential when selecting a brand voice to reflect the desired personality remains. This study addresses this issue and provides three key contributions to the literature on vocal personality perception and voice marketing.

First, a comprehensive and multidisciplinary synthesis of the existing research on vocal personality perception was provided through a systematic literature review covering nearly 85 years of published research on the topic with 52 articles. In doing so, this study is the first to systematically review vocal personality perception literature. The identified perceptual studies incorporated 444 individual findings on personality perception through voice parameters. The results were summarized into a concept matrix scheme, in which the essential data to extract was categorized according to the study's research questions and framework. Descriptive data on the findings was provided within the macro-level analysis.

Second, managerial recommendations were presented on selecting a voice that fits a brand personality based on insights about human personality. For this purpose, findings on human personality perception were structured according to the five brand dimensions of Aaker (1997), sincerity, excitement, competence, sophistication, and ruggedness, and analyzed on a micro level. Vocal profiles for these brand personality dimensions were derived to guide brand voice selection. The scope of this analysis was limited to research that used external judgments to assess perceptions of personality and is conducted in American English.

For the vocal perception of **sincerity**, findings on related personality traits were used to derive managerial recommendations for selecting the appropriate brand voice. Based on the vocal perception of agreeableness, benevolence, credibility, politeness, sensitivity, and social attractiveness, it was concluded that brand personality sincerity is perceived when *male voices* with a high pitch variability or low loudness are used. Various voice qualities like sharpness, thinness, tension, nasality, or flatness mitigate the perception of sincerity-related personality traits. Brands with *female voices* should use a voice with high pitch variability, breathiness, tension, or less resonance to communicate sincerity. Furthermore, a medium speaking rate is preferable, and vocal qualities such as nasality or roughness should be avoided, as they are perceived to decrease politeness.

Based on the vocal perception of activity, boldness, enthusiasm, extraversion, talkativeness, youthfulness, and young, it was concluded that the brand personality **excitement** is perceived when *male voices* with a high pitch, high loudness, or a dynamic intonation are used. Additionally, the voice should be resonant or breathy but less rough or tense to be assessed with excitement. Concerning the speaking rate and nasality of male voices, marketers should consider a tradeoff; depending on which personality trait is to be communicated, the specifications of these voice parameters must be adjusted. An increased speaking rate leads to the perception of activity, enthusiasm, extraversion, and talkativeness but decreases the perception of boldness; nasality seems to reduce the perception of activity and enthusiasm but increases the perception of extraversion. Exciting brands with *female voices* should be loud or have a dynamic intonation. Additionally, an increased speaking rate, breathiness, tension, or resonance improves the perception of excitement. Voices with roughness, nasality, or creakiness should be avoided as they decrease assessments of excitement-related personality traits like activity, enthusiasm, talkativeness, and youthfulness. Further, research findings indicated that exciting brands that sound young should have less resonant female voices, as resonance can indicate maturity.

A **competent** brand is best perceived through a *male voice* with high loudness, high pitch variability, low pitch, or a fast speaking rate. The linear combination of high pitch variability and a fast speaking rate increases competence perception, whereas a low pitch variability and a slow speaking rate leads to the opposite. *Female voices* should also be high in loudness, high in pitch variability, or should have a fast speaking rate. A creaky voice should be avoided, as creakiness diminishes the perception of competence. Conflicting results existed concerning the competence perception of female pitch, indicating the need for more research on female pitch perception.

Whereas the brand personality dimensions of sincerity, excitement, and competence strongly relate to the Big Five personality dimensions of agreeableness, extraversion, and consciousness, the remaining brand personalities, sophistication, and ruggedness, differ from any Big Five personalities. This distinction is evident in the literature, as these two brand personalities or related personalities had the fewest results, one of the most significant gaps in research. Regarding the perception of **sophistication** and the related personality "feminity", Addington (1968) showed that sophistication is perceived through resonant, rough, or less nasal *male voices*. For *female voices*, results indicated that various voice qualities like breathiness, nasality, roughness, tension, and thinness decrease the perception of sophistication. Concerning the perception of feminity, results for female voices contradict those of sophistication perception since breathiness, tension, and thinness increase feminity assessments. In contrast, rough and flat voices decrease perceptions of feminity. Assessments of feminity increase in male voices with a high pitch variability. Since masculinity (for *male voices*) is the only **ruggedness**-related personality trait found within the extracted data, it was impossible to make recommendations for selecting an appropriate voice for rugged brands.

The study's third contribution is an overview of future research directions for vocal brand and human personality perception. The most significant recommendation was that more research is generally needed into the vocal perception of brand personalities, as literature on vocal personality perception has concentrated on human personalities. In this regard, it would be interesting to find out whether the identified brand personalities can be represented acoustically or whether there are brand personalities that are difficult or even impossible to describe vocally. Besides the general lack of research on the perception of brand personality through voice, five avenues for future brand and human personality research based on identified research gaps were derived.

It was concluded that research on the perception of voice qualities, like breathiness, creakiness, roughness, or hoarseness, is needed since little research exists on this soundwave dimension. Although measuring voice quality parameters remains challenging, this dimension is an essential part of the uniqueness of a voice and, therefore, contributes significantly to personality perception. Moreover, future studies on vocal personality perception should employ a more balanced approach by including voices of both genders, rather than focusing exclusively on either male or female voices. In particular, research on the perception of female voices is currently underrepresented, a gap that must be addressed. Further, research on vocal personality perception in various languages and cultures is needed since there is a surplus of American English studies. Given the impact of cultural norms on the way individuals use their voice, which in turn affects how their personalities are perceived through the voice (Krauss et al., 2002; Waaramaa et al., 2021), further studies conducted in different languages, in addition to cross-cultural studies, are required. The last two research recommendations are methodological. The literature review shows that the relationship between single voice parameters and personalities is primarily examined, even though it is known that voice parameters correlate with each other in natural speech (Apple et al., 1979). Thus, their combined effect on personality perception should be of interest. Therefore, further research on multivariate analyses of voice parameters is recommended, which is likely to be facilitated by current advances in speech processing software and machine learning techniques. Lastly, meta-analyses on the vocal perception of personalities should be performed to provide in-depth, more accurate, and quantitative insights into how voice parameters affect personality perception.

8.1. Limitations

This study was constrained by limitations necessary to narrow the scope of a feasible literature review. For example, studies on the automatic vocal perception of personality through machine learning were excluded, as they might provide additional insights into how voices, in general, are perceived. Moreover, studies in which non-native speakers, children, and the elderly were used as speakers or listeners were not included in speech production, as personality perceptions would be influenced by factors that could not be controlled in a qualitative review, or the scope would be unmanageable.

In addition to the predefined limitations concerning the scope of the study, further restrictions emerged during the analyses and interpretation of the findings. The absence of prior research on the vocal perception of specific brand personalities led to the formulation of recommendations on selecting a brand voice based on findings related to human personalities. Except for the personality dimensions of competence and sophistication, no direct results were obtained for the dimensions of sincerity, excitement, and ruggedness of the BPS of Aaker (1997). The managerial recommendations for selecting a sophisticated and rugged brand voice were limited because the results on sophistication perception were derived from a single study, and masculinity was the only ruggedness-related personality trait identified in the literature on perceptual studies.

Moreover, the micro-level analysis was constrained to studies that employed external judgments to assess personality perception and were conducted in American English. This limitation was because most results included externally assessed personality perceptions in American English, facilitating consolidation and comparison of the findings.

Essay 1 References

Aaker, J. L. (1997). Dimensions of Brand Personality. *Journal of Marketing Research*, 34(3), 347–356. https://doi.org/10.1177/002224379703400304

Aaker, J. L., Benet-Martínez, V., & Garolera, J. (2001). Consumption Symbols as Carriers of Culture: A Study of Japanese and Spanish Brand Personality Constructs. *Journal of Personality and Social Psychology*, 81(3), 492–508. https://doi.org/10.1037/0022-3514.81.3.492

Addington, D. W. (1968). The relationship of selected vocal characteristics to personality perception. *Speech Monographs*, 35(4), 492–503. https://doi.org/10.1080/03637756809375599

Afshan, A., Guo, J., Park, S. J., Ravi, V., Flint, J., & Alwan, A. (2018). Effectiveness of Voice Quality Features in Detecting Depression. *Interspeech* 2018, 1676–1680. https://doi.org/10.21437/Interspeech.2018-1399

Aggarwal, P., & McGill, A. L. (2007). Is that car smiling at me? Schema congruity as a basis for evaluating anthropomorphized products. *Journal of Consumer Research*, 34(4), 468–479. https://doi.org/10.1086/518544

Allport, G. W., & Cantril, H. (1934). Judging personality from voice. *The Journal of Social Psychology*, 5(1), 37–55.

Anderson, R. C., Klofstad, C. A., Mayew, W. J., & Venkatachalam, M. (2014). Vocal fry may undermine the success of young women in the labor market. *PLOS ONE*, 9(5), e97506. https://doi.org/10.1371/journal.pone.0097506

Anjos de Oliveira, A., Dajer, M., Fernandes, P., & Teixeira, J. (2020). Clustering of Voice Pathologies based on Sustained Voice Parameters. In *13th International Joint Conference on Biomedical Engineering Systems and Technologies* (Vol. 4, pp. 280–287). https://doi.org/10.5220/0009146202800287

Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, 37(5), 715–727. https://doi.org/10.1037/0022-3514.37.5.715

Aronovitch, C. D. (1976). The voice of personality: Stereotyped judgments and their relation to voice quality and sex of speaker. *The Journal of Social Psychology*, 99(2), 207–220. https://doi.org/10.1080/00224545.1976.9924774

Atari, M., & Henrich, J. (2023). Historical Psychology. *Current Directions in Psychological Science*, 32(2), 176–183. https://doi.org/10.1177/09637214221149737

Azoulay, A., & Kapferer, J.-N. (2003). Do brand personality scales really measure brand personality? *Journal of Brand Management*, 11(2), 143–155. https://doi.org/10.1057/palgrave.bm.2540162

Bank of America. (n.d.). *Erica is here for you, your life and your goals: Spend, save and plan smarter with Erica, your virtual financial assistant.* https://promotions.bankofamerica.com/digitalbanking/mobilebanking/erica

Barsties V Latoszek, B., Kim, G.-H., Delgado Hernández, J., Hosokawa, K., Englert, M., Neumann, K., & Hetjens, S. (2021). The validity of the Acoustic Breathiness Index in the evaluation of breathy voice quality: A Meta-Analysis. *Clinical Otolaryngology*, 46(1), 31–40. https://doi.org/10.1111/coa.13629

Barsties V Latoszek, B., Lehnert, B., & Janotte, B. (2020). Validation of the Acoustic Voice Quality Index Version 03.01 and Acoustic Breathiness Index in German. *Journal of Voice*, 34(1), 157.e17-157.e25. https://doi.org/10.1016/j.jvoice.2018.07.026

Barsties V Latoszek, B., Mathmann, P., & Neumann, K. (2021). The cepstral spectral index of dysphonia, the acoustic voice quality index and the acoustic breathiness index as novel multiparametric indices for acoustic assessment of voice quality. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 29(6), 451–457. https://doi.org/10.1097/MOO.0000000000000743

Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding voice perception. *British Journal of Psychology*, 102(4), 711–725. https://doi.org/10.1111/j.2044-8295.2011.02041.x

Belin, P., Fecteau, S., & Bédard, C. (2004). Thinking the voice: Neural correlates of voice perception. *Trends in Cognitive Sciences*, 8(3), 129–135. https://doi.org/10.1016/j.tics.2004.01.008

Blasi, D. E., Henrich, J., Adamou, E., Kemmerer, D., & Majid, A. (2022). Over-reliance on English hinders cognitive science. *Trends in Cognitive Sciences*, 26(12), 1153–1170. https://doi.org/10.1016/j.tics.2022.09.015

BMW Group. (2018, September 6). *„Hey BMW, jetzt reden wir." Mit BMWs Intelligent Personal Assistant bekommt der BMW eine Persönlichkeit* [Press release]. https://www.press.bmwgroup.com/deutschland/article/detail/T0284429DE/%E2%80%9Ehey-bmw-jetzt-reden-wir-%E2%80%9C-mit-bmws-intelligent-personal-assistant-bekommt-der-bmw-eine-persoenlichkeit?language=de

Boersma, P., & van Heuven, V. J. (2001). Speak and unSpeak with Praat. *Glot International*, 5(9/10), 341–347.

Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, 82(1), 55–59. https://doi.org/10.1016/j.anbehav.2011.03.024

Bosnjak, M., Bochmann, V., & Hufschmidt, T. (2007). Dimensions of brand personality attributions: A person-centric approach in the German cultural context. *Social Behavior and Personality: An International Journal*, 35(3), 303–316. https://doi.org/10.2224/sbp.2007.35.3.303

Brave, S., & Nass, C. (2009). Emotion in Human-Computer Interaction. In A. Sears & J. Jacko (Eds.), *Human Factors and Ergonomics. Human-Computer Interaction Fundamentals* (pp. 53–68). CRC Press. https://doi.org/10.1201/b10368-6

Brown, B. L., Giles, H., & Thakerar, J. N. (1985). Speaker evaluations as a function of speech rate, accent and context. *Language & Communication*, 5(3), 207–220. https://doi.org/10.1016/0271-5309(85)90011-4

Brown, B. L., Strong, W. J., & Rencher, A. C. (1972). Acoustic Parameters of Personality. *The Journal of the Acoustical Society of America*, 51(1A), 121. https://doi.org/10.1121/1.1981348

Brown, B. L., Strong, W. J., & Rencher, A. C. (1973). Perceptions of personality from speech: Effects of manipulations of acoustical parameters. *The Journal of the Acoustical Society of America*, 54(1), 29–35. https://doi.org/10.1121/1.1913571

Brown, B. L., Strong, W. J., & Rencher, A. C. (1974). Fifty-four voices from two: The effects of simultaneous manipulations of rate, mean fundamental frequency, and variance of fundamental frequency on ratings of personality from speech. *The Journal of the Acoustical Society of America*, 55(2), 313–318. https://doi.org/10.1121/1.1914504

Bryant, G. A. (2022). Vocal communication across cultures: Theoretical and methodological issues. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 377(1841), 20200387. https://doi.org/10.1098/rstb.2020.0387

Caprara, G. V., Barbaranelli, C., & Guido, G. (2001). Brand personality: How to make the metaphor fit? *Journal of Economic Psychology*, 22(3), 377–395. https://doi.org/10.1016/s0167-4870(01)00039-3

Carpenter, C. J. (2012). A Meta-Analysis and an Experiment Investigating the Effects of Speaker Disfluency on Persuasion. *Western Journal of Communication*, 76(5), 552–569. https://doi.org/10.1080/10570314.2012.662307

Caspi, A., & Roberts, B. W. (2001). Personality Development Across the Life Course: The Argument for Change and Continuity. *Psychological Inquiry*, 12(2), 49–66. https://doi.org/10.1207/S15327965PLI1202_01

Chandler, J., & Schwarz, N. (2010). Use does not wear ragged the fabric of friendship: Thinking of objects as alive makes people less willing to replace them. *Journal of Consumer Psychology*, 20(2), 138–145. https://doi.org/10.1016/j.jcps.2009.12.008

Charoenruk, N., & Olson, K. (2018). Do Listeners Perceive Interviewers' Attributes from their Voices and Do Perceptions Differ by Question Type? *Field Methods*, 30(4), 312–328. https://doi.org/10.1177/1525822X18784500

Chen, A., Rietveld, T., & Gussenhoven, C. (2001). Language-specific effects of pitch range on the perception of universal intonational meaning. In *7th European Conference on Speech Communication and Technology (Eurospeech 2001)* (pp. 1403–1406). https://doi.org/10.21437/Eurospeech.2001-362

Cheng, A. (2018, June 20). Amazon-Marriott Deal Will Make Alexa A Hotel Butler, But The Implications Range Far Wider. *Forbes*. https://www.forbes.com/sites/andriacheng/2018/06/19/amazons-marriott-deal-is-way-beyond-alexa-as-your-new-hotel-butler/

Chu, S.-C., & Sung, Y. (2011). Brand personality dimensions in China. *Journal of Marketing Communications*, 17(3), 163–181. https://doi.org/10.1080/13527260903387931

Clark, J., Yallop, C., & Fletcher, J. (2007). *An introduction to phonetics and phonology* (3rd ed.). Wiley-Blackwell.

Collins, S. A. (2000). Men's voices and women's choices. *Animal Behaviour*, 60(6), 773–780. https://doi.org/10.1006/anbe.2000.1523

Collins, S. A., & Missing, C. (2003). Vocal and visual attractiveness are related in women. *Animal Behaviour*, 65(5), 997–1004. https://doi.org/10.1006/anbe.2003.2123

Conrad, F. G., Schober, M. F., & Dijkstra, W. (2008). Cues of Communication Difficulty in Telephone Interviews. In Lepkowski, J. M., Tucker, C., J. M. Brick, E. de Leeuw, L. Japec, P. J. Lavrakes, M. W. L. Link, & R. L. Sangster (Eds.), *Advances in Telephone Survey Methodology* (pp. 212–230). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470173404.ch10

Dahl, D. W. (2010). Understanding the role of spokesperson voice in broadcast advertising. In A. Krishna (Ed.), *Sensory marketing: Research on the Sensuality of Products* (pp. 169–182). Routledge. https://doi.org/10.4324/9780203892060-21 d'Astous, A., & Lévesque, M. (2003). A scale for measuring store personality. *Psychology & Marketing*, 20(5), 455–469. https://doi.org/10.1002/mar.10081

Feinberg, D. R., Debruine, L. M., Jones, B. C., & Perrett, D. I. (2008). The role of femininity and averageness of voice pitch in aesthetic judgments of women's voices. *Perception*, 37(4), 615–623. https://doi.org/10.1068/p5514

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, 69(3), 561–568. https://doi.org/10.1016/j.anbehav.2004.06.012

Ferrandi, J. M., Valette-Florence P., & Fine-Falcy S. (2015). Aaker's Brand Personality Scale in a French Context: A Replication and a Preliminary Test of Its Validity. In *Proceedings of the 2000 Academy of Marketing Science (AMS) Annual Conference* (pp. 7–13). https://doi.org/10.1007/978-3-319-11885-7_2

Fitch, W. T., & Giedd, J. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America*, 106(3), 1511–1522. https://doi.org/10.1121/1.427148

Fraccaro, P. J., O'Connor, J. J., Re, D. E., Jones, B. C., Debruine, L. M., & Feinberg, D. R. (2013). Faking it: Deliberately altered voice pitch and vocal attractiveness. *Animal Behaviour*, 85(1), 127–136. https://doi.org/10.1016/j.anbehav.2012.10.016

Freling, T. H., & Forbes, L. P. (2005). An examination of brand personality through methodological triangulation. *Journal of Brand Management*, 13(2), 148–162. https://doi.org/10.1057/palgrave.bm.2540254

Frühholz, S., & Belin, P. (2019). The Science of Voice Perception. In S. Frühholz & P. Belin (Eds.), *The Oxford Handbook of Voice Perception* (1st ed., pp. 2–14). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198743187.013.1

Geuens, M., Weijters, B., & Wulf, K. de (2009). A new measure of brand personality. *International Journal of Research in Marketing*, 26(2), 97–107. https://doi.org/10.1016/j.ijresmar.2008.12.002

Gilmore, G. W. (1919). *Animism: or, thought currents of primitive peoples*. Marshall Jones Company.

Gobl, C., & Chasaide, A. N. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, 40(1-2), 189–212. https://doi.org/10.1016/S0167-6393(02)00082-1

Gocsál, Á. (2009). Female listeners' personality attributions to male speakers: The role of acoustic parameters of speech. *Pollack Periodica*, 4(3), 155–165. https://doi.org/10.1556/Pollack.4.2009.3.14

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *The American Psychologist,* 48(1), 26–34. https://doi.org/10.1037/0003-066x.48.1.26

Gordon, M., & Ladefoged, P. (2001). Phonation types: a cross-linguistic overview. *Journal of Phonetics*, 29(4), 383–406. https://doi.org/10.1006/jpho.2001.0147

Grohmann, B. (2009). Gender Dimensions of Brand Personality. *Journal of Marketing Research*, 46(1), 105–119. https://doi.org/10.1509/jmkr.46.1.105

Guyer, J. J., Fabrigar, L. R., & Vaughan-Johnston, T. I. (2019). Speech Rate, Intonation, and Pitch: Investigating the Bias and Cue Effects of Vocal Confidence on Persuasion. *Personality and Social Psychology Bulletin*, 45(3), 389–405. https://doi.org/10.1177/0146167218787805

Hildebrand, C., Efthymiou, F., Busquet, F., Hampton, W. H., Hoffman, D. L., & Novak, T. P. (2020). Voice analytics in business research: Conceptual foundations, acoustic feature extraction, and applications. *Journal of Business Research*, 121, 364–374. https://doi.org/10.1016/j.jbusres.2020.09.020

Hodges-Simeon, C. R., Gaulin, S. J. C., & Puts, D. A. (2010). Different Vocal Parameters Predict Perceptions of Dominance and Attractiveness. *Human Nature*, 21(4), 406–427. https://doi.org/10.1007/s12110-010-9101-5

Huang, D. Y., Ong, E. P., Rahardja, S., Dong, M., & Li, H. (2009). Transformation of vocal characteristics: A review of literature. *International Scholarly and Scientific Research and Innovation*, 3(12), 2288-2296.

Imhof, M. (2010). Listening to Voices and Judging People. *International Journal of Listening*, 24(1), 19–33. https://doi.org/10.1080/10904010903466295

Inam, I. A., Azeta, A. A., & Daramola, O. (2017). Comparative analysis and review of interactive voice response systems. In *2017 Conference on Information Communication Technology and Society (ICTAS)* (pp. 1–6). IEEE. https://doi.org/10.1109/ictas.2017.7920660

Johnstone, t., & Scherer, K. R. (2000). Vocal communication of emotion. In Lewis M. & Haviland J. (Eds.), *Handbook of Emotion* (2nd ed.). New York: Guilford.

Jones, B. C., Feinberg, D. R., Debruine, L. M., Little, A. C., & Vukovic, J. (2010). A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. *Animal Behaviour*, 79(1), 57–62. https://doi.org/10.1016/j.anbehav.2009.10.003

Jurafsky, D., & Martin, J. H. (2020). *Speech and language processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (3rd ed.).

Ketrow, S. M. (1990). Attributes of a telemarketer's voice and persuasiveness.: A review and synthesis of the literature. *Journal of Direct Marketing*, 4(3), 7–21. https://doi.org/10.1002/dir.4000040304

Kimble, C. E., & Seidel, S. D. (1991). Vocal signs of confidence. *Journal of Nonverbal Behavior*, 15(2), 99–105. https://doi.org/10.1007/BF00998265

Kitchenham, B. (2004). *Procedures for performing systematic reviews*. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=29890a936639862f45cb9a987dd599dce9759bf5

Klasmeyer, G., & Sendlmeier, W. F. (1997). The classification of different phonation types in emotional and neutral speech. *International Journal of Speech, Language and the Law*, 4(1), 104–124. https://doi.org/10.1558/ijsll.v4i1.104

Klofstad, C. A., Anderson, R. C., & Nowicki, S. (2015). Perceptions of Competence, Strength, and Age Influence Voters to Select Leaders with Lower-Pitched Voices. *PLOS ONE*, 10(8), e0133779. https://doi.org/10.1371/journal.pone.0133779

Koutsombogera, M., Sarthy, P., & Vogel, C. (2020). Acoustic Features in Dialogue Dominate Accurate Personality Trait Classification. In *2020 IEEE International Conference on Human-Machine Systems (ICHMS)* (pp. 1–3). https://doi.org/10.1109/ichms49158.2020.9209445

Koutsoumpis, A., & Vries, R. E. de (2022). What Does Your Voice Reveal About You? Trait Activation of Voice Characteristics and Their Relation With Personality and Communication Styles. *Journal of Individual Differences*, 43(3), 160–167. https://doi.org/10.1027/1614-0001/a000362

Krahé, B., & Papakonstantinou, L. (2020). Speaking like a Man: Women's Pitch as a Cue for Gender Stereotyping. *Sex Roles*, 82(1-2), 94–101. https://doi.org/10.1007/s11199-019-01041-z

Krahé, B., Uhlmann, A., & Herzberg, M. (2021). The Voice Gives It Away: Male and Female Pitch as a Cue for Gender Stereotyping. *Social Psychology*, 52(2), 101–113. https://doi.org/10.1027/1864-9335/a000441

Kramer, E. (1964). Personality Stereotypes in Voice: A Reconsideration of the Data. The *Journal of Social Psychology*, 62(2), 247–251. https://doi.org/10.1080/00224545.1964.9919523

Krauss, R. M., Freyberg, R., & Morsella, E. (2002). Inferring speakers' physical attributes from their voices. *Journal of Experimental Social Psychology*, 38(6), 618–625. https://doi.org/10.1016/S0022-1031(02)00510-3

Kreiman, J., & Sidtis, D. (2011). Perception of Emotion and Personality from Voice. In *Foundations of Voice Studies* (pp. 302–360). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444395068.ch9

Kreutzer, R. T., & Vousoghi, D. S. (2020). *Voice-Marketing: Der Siegeszug der digitalen Assistenten.* Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-29474-8_2

Kuang, J., & Liberman, M. (2018). Integrating Voice Quality Cues in the Pitch Perception of Speech and Non-speech Utterances. *Frontiers in Psychology*, 9, 2147. https://doi.org/10.3389/fpsyg.2018.02147

Kumar, A. (2018). Story of Aaker's brand personality scale criticism. *Spanish Journal of Marketing - ESIC*, 22(2), 203–230. https://doi.org/10.1108/SJME-03-2018-005

Kumar, V., Dixit, A., Javalgi, R. G., & Dass, M. (2016). Research framework, strategies, and applications of intelligent agent technologies (IATs) in marketing. *Journal of the Academy of Marketing Science*, 44(1), 24–45. https://doi.org/10.1007/s11747-015-0426-9

Lee, H., & Cho, C.-H. (2017). An application of brand personality to advergames: The effect of company attributes on advergame personality. *Computers in Human Behavior*, 69, 235–245. https://doi.org/10.1016/j.chb.2016.12.035

Lee, H. O., & Boster, F. J. (1992). Collectivism-Individualism in Perceptions of Speech Rate. *Journal of Cross-Cultural Psychology*, 23(3), 377–388. https://doi.org/10.1177/0022022192233008

The Lens. (2023). *What is The Lens*. https://about.lens.org/what/

Levitt, A., & Lucas, M. (2018). Effects of Four Voice Qualities and Formant Dispersion on Perception of a Female Voice. *Psychology of Language and Communication*, 22(1), 394–416. https://doi.org/10.2478/plc-2018-0018

Linnenluecke, M. K., Marrone, M., & Singh, A. K. (2020). Conducting systematic literature reviews and bibliometric analyses. *Australian Journal of Management*, 45(2), 175–194. https://doi.org/10.1177/0312896219877678

Lowe, M. L., & Haws, K. L. (2017). Sounds Big: The Effects of Acoustic Pitch on Product Perceptions. *Journal of Marketing Research*, 54(2), 331–346. https://doi.org/10.1509/jmr.14.0300

Mairesse, F., Walker, M. A., Mehl, M. R., & Moore, R. K. (2007). Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text. *Journal of Artificial Intelligence Research*, 30, 457–500. https://doi.org/10.1613/jair.2349

Mallory, E. B., & Miller, V. R. (1958). A possible basis for the association of voice characteristics and personality traits. *Speech Monographs*, 25(4), 255–260. https://doi.org/10.1080/03637755809375240

Mari, A., Mandelli, A., & Algesheimer, R. (2020). The Evolution of Marketing in the Context of Voice Commerce: A Managerial Perspective. In F. F.-H. Nah & K. Siau (Eds.), *Lecture Notes in Computer Science. HCI in Business, Government and Organizations* (Vol. 12204, pp. 405–425). Springer International Publishing. https://doi.org/10.1007/978-3-030-50341-3_32

Mayew, W. J., Parsons, C. A., & Venkatachalam, M. (2013). Voice pitch and the labor market success of male chief executive officers. Evolution and Human Behavior, 34(4), 243–248. https://doi.org/10.1016/j.evolhumbehav.2013.03.001

McAleer, P., & Belin, P. (2019). The Perception of Personality Traits from Voices. In S. Frühholz & P. Belin (Eds.), *The Oxford Handbook of Voice Perception* (1st ed., pp. 585–606). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198743187.013.26

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say ‘Hello’? Personality Impressions from Brief Novel Voices. *PLOS ONE*, 9(3), 1–9. https://doi.org/10.1371/journal.pone.0090779.t001

Medical Xpress (2024, June 21). Our brains respond differently to human and AI-generated speech, but we still struggle to tell them apart. *Medical Xpress*. https://medicalxpress.com/news/2024-06-brains-differently-human-ai-generated.html

Mehrabian, A., & Ferris, S. R. (1967). Inference of attitudes from nonverbal communication in two channels. *Journal of Consulting Psychology*, 31(3), 248–252. https://doi.org/10.1037/h0024648

Mercedes-Benz Group AG. (2022, April 19). *MBUX reaches a new level* [Press release]. https://group.mercedes-benz.com/innovation/digitalisation/connectivity/mbux-interior-assist.html

Michalsky, J., Niebuhr, O., & Penke, L. (2020). Do charismatic people produce charismatic speech? On the relationship between the Big Five personality traits and prosodic features of speaker charisma in female speakers. In *10th International Conference on Speech* (pp. 700–704). https://doi.org/10.21437/SpeechProsody.2020-143

Miller, N., Maruyama, G., Beaber, R. J., & Valone, K. (1976). Speed of speech and persuasion. *Journal of Personality and Social Psychology*, 34(4), 615–624. https://doi.org/10.1037/0022-3514.34.4.615

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, 8(5), 336-341. https://doi.org/10.1016/j.ijsu.2010.02.007

Moore, W. E. (1939). Personality Traits and Voice Quality Deficiencies. *Journal of Speech Disorders*, 4(1), 33–36. https://doi.org/10.1044/jshd.0401.33

O'Connor, J. J. M., & Barclay, P. (2018). High voice pitch mitigates the aversiveness of anti-social cues in men's speech. *British Journal of Psychology*, 109(4), 812–829. https://doi.org/10.1111/bjop.12310

Oksenberg, L., Coleman, L., & Cannell, C. F. (1986). Interviewers' Voices and Refusal Rates in Telephone Surveys. *Public Opinion Quarterly*, 50(1), 97. https://doi.org/10.1086/268962

Oleszkiewicz, A., Pisanski, K., Lachowicz-Tabaczek, K., & Sorokowska, A. (2017). Voice-based assessments of trustworthiness, competence, and warmth in blind and sighted adults. *Psychonomic Bulletin & Review*, 24(3), 856–862. https://doi.org/10.3758/s13423-016-1146-y

Packard, G., & Berger, J. (2024). The Emergence and Evolution of Consumer Language Research. *Journal of Consumer Research*, 51(1), 42–51. https://doi.org/10.1093/jcr/ucad013

Paluch, S., & Wittkop, T. (2020). Voice Marketing – Die Stimme der Zukunft? In M. Bruhn, C. Burmann, & M. Kirchgeorg (Eds.), *Marketing Weiterdenken: Zukunftspfade für eine marktorientierte Unternehmensführung* (2nd ed. 2020, pp. 509–520). Springer Fachmedien Wiesbaden; Imprint: Springer Gabler. https://doi.org/10.1007/978-3-658-31563-4_26

Paunonen, S. V., Jackson, D. N., Trzebinski, J., & Forsterling, F. (1992). Personality structure across cultures: A multimethod evaluation. *Journal of Personality and Social Psychology,* 62(3), 447–456. https://doi.org/10.1037/0022-3514.62.3.447

Pear, T. H. (1931). *Voice and Personality* (1st ed.). Chapman and Hall, Ltd. https://doi.org/10.1038/128657a0

Pearsell, S., & Pape, D. (2023). The effects of different voice qualities on the perceived personality of a speaker. *Frontiers in Communication*, 7. https://doi.org/10.3389/fcomm.2022.909427

Peng, Y., Zebrowitz, L. A., & Lee, H. K. (1993). The Impact of Cultural Background and Cross-Cultural Experience on Impressions of American and Korean Male Speakers. *Journal of Cross-Cultural Psychology*, 24(2), 203–220. https://doi.org/10.1177/0022022193242005

Pisanski, K., Fraccaro, P. J., Tigue, C. C., O'Connor, J. J., Röder, S., Andrews, P. W., Fink, B., Debruine, L. M., Jones, B. C., & Feinberg, D. R. (2014). Vocal indicators of body size in men and women: a meta-analysis. *Animal Behaviour*, 95, 89–99. https://doi.org/10.1016/j.anbehav.2014.06.011

Pisanski, K., & Rendall, D. (2011). The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. The *Journal of the Acoustical Society of America*, 129(4), 2201–2212. https://doi.org/10.1121/1.3552866

Pompino-Marschall, B. (2009). *Einführung in die Phonetik* (3rd ed.). Walter de Gruyter https://doi.org/10.1515/9783110224818

Poon, M. M., Chan, K. M., & Yiu, E. M. (2018). The Relationship Between Speech Rate, Voice Quality and Listeners’ Purchase Intentions. In *9th International Conference on Speech Prosody 2018* (pp. 468–472). https://doi.org/10.21437/speechprosody.2018-95

Puts, D. A., Barndt, J. L., Welling, L. L., Dawood, K., & Burriss, R. P. (2011). Intrasexual competition among women: Vocal femininity affects perceptions of attractiveness and flirtatiousness. *Personality and Individual Differences*, 50(1), 111–115. https://doi.org/10.1016/j.paid.2010.09.011

Puts, D. A., Gaulin, S. J., & Verdolini, K. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior*, 27(4), 283–296. https://doi.org/10.1016/j.evolhumbehav.2005.11.003

Puts, D. A., Hodges, C. R., Cárdenas, R. A., & Gaulin, S. J. (2007). Men's voices as dominance signals: vocal fundamental and formant frequencies influence dominance attributions among men. *Evolution and Human Behavior*, 28(5), 340–344. https://doi.org/10.1016/j.evolhumbehav.2007.05.002

Rauschnabel, P. A., & Ahuvia, A. C. (2014). You’re so lovable: Anthropomorphism and brand love. *Journal of Brand Management*, 21, 372–395. https://doi.org/10.1057/bm.2014.14

Ray, G. B. (1986). Vocally cued personality prototypes: An implicit personality theory approach. *Communication Monographs*, 53(3), 266–276. https://doi.org/10.1080/03637758609376141

Ray, G. B., Ray, E. B., & Zahn, C. J. (1991). Speech behavior and social evaluation: An examination of medical messages. *Communication Quarterly*, 39(2), 119–129. https://doi.org/10.1080/01463379109369790

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people*. Cambridge University Press.

Riding, D., Lonsdale, D., & Brown, B. (2006). The Effects of Average Fundamental Frequency and Variance of Fundamental Frequency on Male Vocal Attractiveness to Women. *Journal of Nonverbal Behavior*, 30(2), 55–61. https://doi.org/10.1007/s10919-006-0005-3

Rodero, E. (2012). A comparative analysis of speech rate and perception in radio bulletins. Text & Talk, 32(3), 391–411. https://doi.org/10.1515/text-2012-0019

Rodero, E. (2013). The Perception of a Broadcasting Voice. *US-China Education Review*, 3(4), 225–230.

Rosenberg, A., & Hirschberg, J. (2009). Charisma perception from text and speech. *Speech Communication,* 51(7), 640–655. https://doi.org/10.1016/j.specom.2008.11.001

Scherer, K. R. (1972). Judging personality from voice: A cross-cultural approach to an old issue in interpersonal perception. *Journal of Personality*, 40(2), 191–210. https://doi.org/10.1111/j.1467-6494.1972.tb00998.x

Scherer, K. R. (1974). Voice quality analysis of American and German speakers. *Journal of Psycholinguistic Research,* 3(3), 281–298. https://doi.org/10.1007/BF01069244

Scherer, K. R. (1978). Personality inference from voice quality: The loud voice of extroversion. European *Journal of Social Psychology*, 8(4), 467–487. https://doi.org/10.1002/ejsp.2420080405

Sheikh, H., Prins, C., & Schrijvers, E. (2023). Artificial Intelligence: Definition and Background. *Mission AI*, 15–41. https://doi.org/10.1007/978-3-031-21448-6_2

Signorello, R., Derrico, F., Poggi, I., & Demolin, D. (2012). How Charisma Is Perceived from Speech: A Multidimensional Approach. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing* (pp. 435–440). https://doi.org/10.1109/socialcom-passat.2012.68

Skjegstad, C. L., & Frühholz, S. (2024). Neural dynamics of processing natural and digital emotions [Poster Session]. In *FENS 2024 Conference Programm*. https://fens2024.abstractserver.com/program/#/details/presentations/4774

Skoog Waller, S., & Eriksson, M. (2016). Vocal Age Disguise: The Role of Fundamental Frequency and Speech Rate and Its Perceived Effects. *Frontiers in Psychology*, 7. https://doi.org/10.3389/fpsyg.2016.01814

Smith, B. L., Brown, B. L., Strong, W. J., & Rencher, A. C. (1975). Effects of speech rate on personality perception. *Language and Speech*, 18(2), 145–152. https://doi.org/10.1177/002383097501800203

Smith, S. M., & Shaffer, D. R. (1991). Celerity and Cajolery: Rapid Speech May Promote or Inhibit Persuasion through its Impact on Message Elaboration. *Personality and Social Psychology Bulletin,* 17(6), 663–669. https://doi.org/10.1177/0146167291176009

Steiner, P. (2018). *Sound Branding: Grundlagen akustischer Markenführung* (3rd ed.). Springer Fachmedien Wiesbaden. https://link.springer.com/content/pdf/10.1007%2F978-3-658-22638-1.pdf https://doi.org/10.1007/978-3-658-22638-1

Stern, J., Schild, C., Jones, B. C., DeBruine, L. M., Hahn, A., Puts, D. A., Zettler, I., Kordsmeyer, T. L., Feinberg, D. R., Zamfir, D., Penke, L., & Arslan, R. C. (2021). Do voices carry valid information about a speaker's personality? *Journal of Research in Personality*, 92, 104092. https://doi.org/10.1016/j.jrp.2021.104092

Stewart, M. A., & Ryan, E. B. (1982). Attitudes Toward Younger and Older Adult Speakers: Effects of Varying Speech Rates. *Journal of Language and Social Psychology*, 1(2), 91–109. https://doi.org/10.1177/0261927X8200100201

Street, R. L., & Brady, R. M. (1982). Speech rate acceptance ranges as a function of evaluative domain, listener speech rate, and communication context. *Communication Monographs*, 49(4), 290–308. https://doi.org/10.1080/03637758209376091

Street, R. L., Brady, R. M., & Putman, W. B. (1983). The Influence of Speech Rate Stereotypes and Rate Similarity or Listeners' Evaluations of Speakers. *Journal of Language and Social Psychology*, 2(1), 37–56. https://doi.org/10.1177/0261927X8300200103

Sung, Y., & Tinkham, S. F. (2005). Brand Personality Structures in the United States and Korea: Common and Culture-Specific Factors. *Journal of Consumer Psychology*, 15(4), 334–350. https://doi.org/10.1207/s15327663jcp1504_8

Tigue, C. C., Borak, D. J., O'Connor, J. J., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, 33(3), 210–216. https://doi.org/10.1016/j.evolhumbehav.2011.09.004

Trouvain, J., Weiss, B., & Barkat-Defradas, M. (2021). Voice Attractiveness: Concepts, Methods, and Data. In *Voice attractiveness: Studies on sexy, likable, and charismatic speakers* (pp. 3–16). Springer, Singapore. https://doi.org/10.1007/978-981-15-6627-1_1

Tsantani, M. S., Belin, P., Paterson, H. M., & McAleer, P. (2016). Low Vocal Pitch Preference Drives First Impressions Irrespective of Context in Male Voices but Not in Female Voices. *Perception*, 45(8), 946–963. https://doi.org/10.1177/0301006616643675

Tusing, K. J., & Dillard, J. P. (2000). The sounds of dominance. *Human Communication Research,* 26(1), 148–171. https://doi.org/10.1111/j.1468-2958.2000.tb00754.x

Usakli, A., & Baloglu, S. (2011). Brand personality of tourist destinations: An application of self-congruity theory. *Tourism Management*, 32(1), 114–127. https://doi.org/10.1016/j.tourman.2010.06.006

van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, 38(3), 253–265. https://doi.org/10.1177/002383099503800303

van Borsel, J., Janssens, J., & Bodt de, M. (2009). Breathiness as a feminine voice characteristic: A perceptual approach. *Journal of Voice: Official Journal of the Voice Foundation*, 23(3), 291–294. https://doi.org/10.1016/j.jvoice.2007.08.002

van Mersbergen, M. (2011). Voice Disorders and Personality: Understanding Their Interactions. *Perspectives on Voice and Voice Disorders*, 21(1), 31–38. https://doi.org/10.1044/vvd21.1.31

voicebot.ai, & Business Wire. (2020). *Number of digital voice assistants in use worldwide from 2019 to 2024 (in billions).* Statista. https://www.statista.com/statistics/973815/worldwide-digital-voice-assistant-in-use/

Waaramaa, T., Lukkarila, P., Järvinen, K., Geneid, A., & Laukkanen, A.-M. (2021). Impressions of Personality from Intentional Voice Quality in Arabic-Speaking and Native Finnish-Speaking Listeners. *Journal of Voice: Official Journal of the Voice Foundation,* 35(2), 326.e21-326.e28. https://doi.org/10.1016/j.jvoice.2019.09.003

Wang, L., Huang, N., Hong, Y., Liu, L., Guo, X., & Chen, G. (2023). Voice-based AI in call center customer service: A natural field experiment. *Production and Operations Management,* 32(4), 1002–1018. https://doi.org/10.1111/poms.13953

Wee, T. T. T. (2004). Extending human personality to brands: The stability factor. *Journal of Brand Management,* 11(4), 317–330. https://doi.org/10.1057/palgrave.bm.2540176

Weiss, B., & Burkhardt, F. (2010). Voice Attributes Affecting Likability Perception. In *Interspeech 2010* (pp. 2014-2017). https://doi.org/10.21437/interspeech.2010-570

Weiss, B., & Burkhardt, F. (2012). Is 'not bad' good enough? Aspects of unknown voices' likability. In *Interspeech 2012* (pp. 510–513). https://doi.org/10.21437/interspeech.2012-97

Wiener, H. J. D., & Chartrand, T. L. (2014). The Effect of Voice Quality on Ad Efficacy. *Psychology & Marketing*, 31(7), 509–517. https://doi.org/10.1002/mar.20712

Woodall, W. G., & Burgoon, J. K. (1984). Talking fast and changing attitudes: A critique and clarification. *Journal of Nonverbal Behavior*, 8(2), 126–142. https://doi.org/10.1007/BF00986999

Wu, Q., Liu, Y., Li, D., Leng, H., Iqbal, Z., & Jiang, Z. (2021). Understanding one's character through the voice: Dimensions of personality perception from Chinese greeting word "Ni Hao". *The Journal of Social Psychology*, 161(6), 653–663. https://doi.org/10.1080/00224545.2020.1856026

Yuasa, I. P. (2010). Creaky voice: A new feminine voice quality for young urban-oriented upwardly mobile American women? *American Speech*, 85(3), 315–337. https://doi.org/10.1215/00031283-2010-018

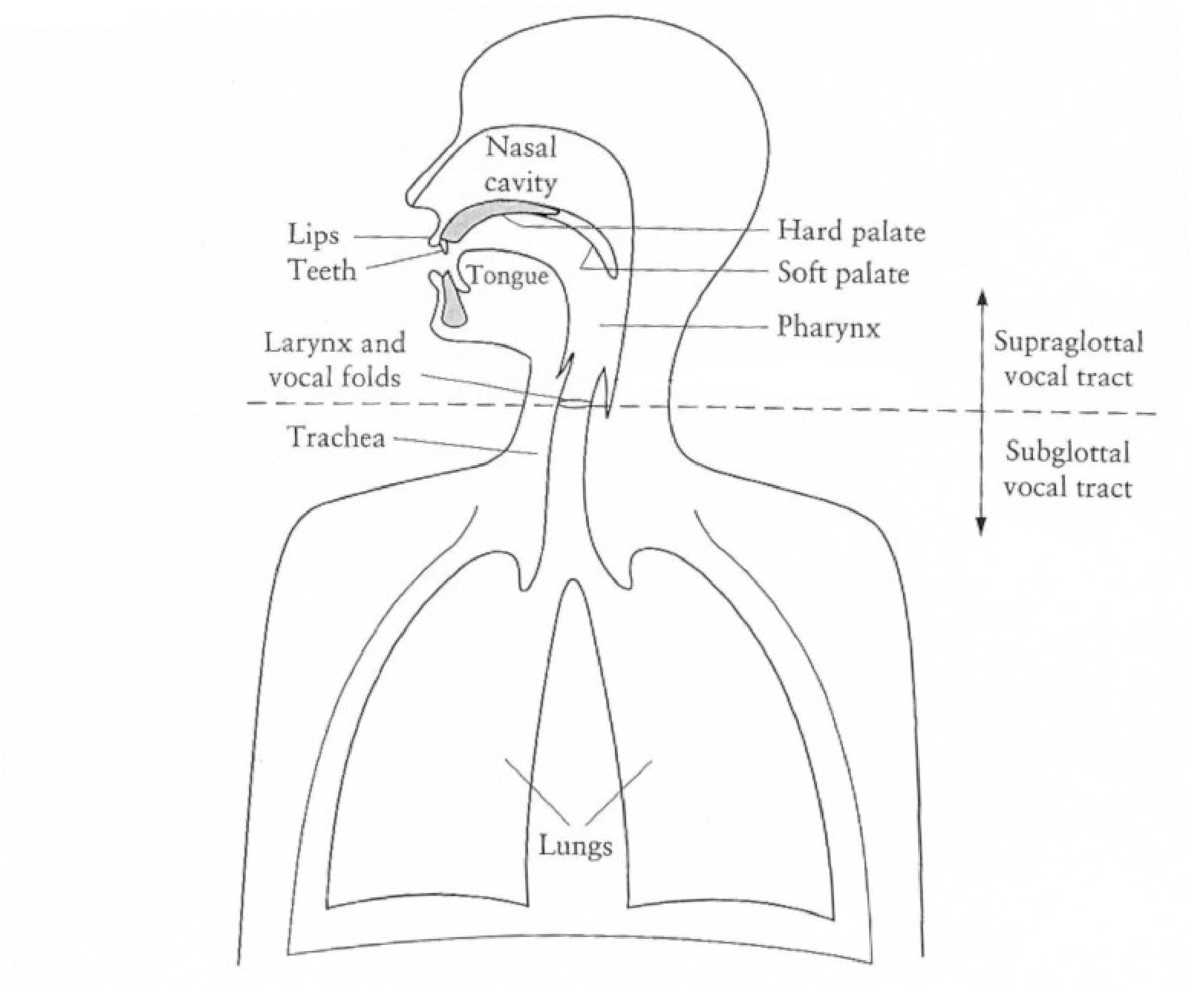
Zäske, R., Skuk, V. G., Golle, J., & Schweinberger, S. R. (2020). The Jena Speaker Set (JESS)-A database of voice stimuli from unfamiliar young and old adult speakers. *Behavior Research Methods*, 52(3), 990–1007. https://doi.org/10.3758/s13428-019-01296-0

Zoghaib, A. (2017). The contribution of a brand spokesperson's voice to consumer-based brand equity. *Journal of Product & Brand Management*, 26(5), 492–502. https://doi.org/10.1108/JPBM-06-2016-1230

Zuckerman, M., & Miyake, K. (1993). The attractive voice: What makes it so? *Journal of Nonverbal Behavior,* 17(2), 119–135. https://doi.org/10.1007/bf01001960

Essay 1 Appendix

Appendix A. Schematic Diagrams of the Organs of Speech Production



Source: Clark et al. (2007), p. 1

Appendix B. List of Excluding Search Terms

Reason for exclusion	Excluding search term
Reference to voice and voice aspects in non-related domains	phoneme, noise, singing, voip-speech-in-noise, music, on-set, actor, actress, color, colour, racial, black, discrimination, “hate speech”, accent, dialect
Reference to visual stimuli for personality perception in combination with voice	fac*, *visual, multisensory
Reference to machine learning and voice/speech recognition	neural, “deep learning”, “voice recognition”, “speech recognition”, TTS
Reference to a medical context	cochlear, Parkinson, patient*, disease*, disabled, disorder*, patholog*, clinic*, therapy, asthma, surgery, blind, dysarthri*, aid, autism, autistic, loss, schizophreni*, cortex, depression
Reference to non-native speakers	EFL, multiling*, “second language”, foreign, L2
Reference to studies with animals, children, or elderly people	aging, “older listener”, child*, old*, age, young*, animal

Notes: The asterisk (*) is a wildcard that searches for any word ending or beginning (e.g., fac* = face, facial). The quotation marks (“ ”) are used for the search of phrases or words that occur in a specific order (e.g., “second language”).
Abbreviations: EFL = English as a Foreign Language; L2 = second language; TTS = text-to-speech.

Appendix C. Detailed Search Settings per Database

Database	Search settings
Scopus	Basic search: article title Language: English Refine results: Limit to document type “research article” Exclusion of excluding search terms with operator “AND NOT”
APA PsycNet	Advanced Search: search terms in “title” Language: English All years Checked boxes: peer-reviewed
Web of Science	Web of Science Core Collection All Editions Document types: article Language: English
ProQuest	Advanced Search: search terms in document title Excluding search terms in abstract and summary text Publication date: all dates Language: English Source types excluded: all source types excluded except scholarly journals

Appendix D. Personality Perception of Voice Parameters of the Timing Dimension

Author	Voice Parameter	Personality	Relationship	Gender of Speaker	Type of Speech Data	Language	Study Design	Personality Assessment Method
Addington (1968)	Fluency (Speaking rate)	Activity/ Energy	positive	male	Read text	American English	Cue manipulation	External judgment
		Activity/ Energy	positive	female				
		Artistic	negative	female				
		Carefulness	negative	female				
		Enthusiasm	positive	female				
		Enthusiasm	positive	male				
		Humor	positive	male				
		Idealistic	negative	female				
		Neuroticism	positive	female				
		Talkativeness	positive	female				
		Talkativeness	positive	male				
		Willingness to Cooperate	positive	female				
Apple et al. (1979)	Fluency (Speaking rate)	Credibility	inverted U-shaped relationship	male	Spontaneous speech	American English	Cue synthesis	External judgment
		Empathy	inverted U-shaped relationship					
		Nervousness	U-shaped relationship					
		Passivity	negative					
		Persuasion	inverted U-shaped relationship					
Aronovitch (1976)	Fluency (Speaking rate)	Activity/ Energy	positive	female	Spontaneous speech	American English	Correlation study	External judgment
		Boldness	positive					
		Social Dominance	positive					
		Emotionality	negative					
		Extraversion	positive					
		Self-Confidence	positive					
		Boldness	negative	male				

		Self-Confidence	positive					
Brown et al. (1985)	Fluency (Speaking rate)	Activity/ Energy	positive	male	Read text	British English	Cue manipulation	External judgment
		Ambition	positive					
		Dependability	negative					
		Intelligence	positive					
		Kindness	negative					
		Politeness	negative					
Brown et al. (1972)	Fluency (Speaking rate)	Benevolence	negative	male	Read text	American English	Cue synthesis	External judgment
		Competence	positive					
Brown et al. (1973)	Fluency (Speaking rate)	Benevolence	negative	male	Read text	American English	Cue synthesis	External judgment
		Competence	positive					
Brown et al. (1974)	Fluency (Speaking rate)	Benevolence	negative	male	Read text	American English	Cue synthesis	External judgment
		Competence	positive					
Charoenruk and Olson (2018)	Fluency (Voice breaks)	Confidence	negative	male	Read text	American English	Correlation study	External judgment
		Confidence	negative	female				
		Reliability	negative	male				
		Reliability	negative	female				
		Trustworthiness	negative	female				
	Fluency (Speaking rate)	Confidence	positive	female				
		Confidence	positive	male				
		Reliability	positive	female				
		Reliability	positive	male				
		Trustworthiness	positive	female				
		Trustworthiness	positive	male				
Gocsál (2009)	Fluency (Speaking rate)	Extraversion	positive	male	Spontaneous speech	Hungarian	Correlation study	External judgment
		Openness to Experience	positive					
	Fluency (Length of pauses)	Extraversion	negative					
		Agreeableness	negative					
		Openness to Experience						

Guyer et al. (2019)	Fluency (Speaking rate)	Confidence	positive	female	Read text	Canadian English	Cue synthesis	External judgment
Kout-sombogera et al. (2020)	Fluency (Voice breaks)	Openness to Experience	positive	female	Read text & Vowels	Dutch	Correlation study	Self-ratings
		Openness to Experience	positive	male				
	Fluency (Speaking rate)	Emotionality	negative	female				
		Emotionality	negative	male				
H. O. Lee and Boster (1992)	Fluency (Speaking rate)	Social Attractiveness	inverted U-shaped relationship	female	Read text	American English	Cue synthesis	External judgment
		Social Attractiveness	positive	male		American English		
		Social Attractiveness	negative	male		Korean		
		Competence	positive	male		American English		
		Competence	positive	female		Korean		
		Competence	negative	male		Korean		
		Trustworthiness	positive	male		American English		
		Trustworthiness	negative	male		Korean		
Mallory and Miller (1958)	Fluency (Speaking rate)	Submissiveness	positive	female	Spontaneous speech	American English	Correlation study	Self-ratings
Miller et al. (1976)	Fluency (Speaking rate)	Intelligence	positive	unknown (probably male)	Read text	American English	Cue manipulation	External judgment
		Knowledgeability	positive	unknown (probably male)				
		Persuasion	positive	unknown (probably male)				
Oksenberg et al. (1986)	Fluency (Speaking rate)	Competence	positive	female	Semi-spontaneous speech	American English	Correlation study	External judgment
Peng et al. (1993)	Fluency (Speaking rate)	Competence	positive	male	Read text	American English	Correlation study	External judgment
Ray et al. (1991)	Fluency (Speaking rate)	Social Attractiveness	inverted U-shaped relationship	female	Read text	American English	Cue manipulation	External judgment
		Dynamic	positive	female				
Ray (1986)	Fluency (Speaking rate)	Benevolence	negative	male	Read text	American English	Cue manipulation	External judgment
		Competence	positive					
B. L. Smith et al. (1975)	Fluency (Speaking rate)	Benevolence	inverted U-shaped relationship	male	Read text	American English	Cue synthesis	External judgment
		Competence	positive					
	Fluency	Credibility	positive	male	Read text	American English	Cue synthesis	External judgment

S. M. Smith and Shaffer (1991)	(Speaking rate)	Credibility	inverted U-shaped relationship					
Street and Brady (1982)	Fluency (Speaking rate)	Social Attractiveness	positive	male	Spontaneous speech	American English	Cue manipulation	External judgment
		Competence	positive					
Street et al. (1983)	Fluency (Speaking rate)	Social Attractiveness	positive	male	Read text	American English	Cue manipulation	External judgment
		Competence	positive					
Tusing and Dillard (2000)	Fluency (Speaking rate)	Social Dominance	negative	female	Semi-spontaneous speech	American English	Correlation study	External judgment
		Social Dominance	negative	male				
Woodall and Burgoon (1984)	Fluency (Speaking rate)	Composure	negative	male	Read text	American English	Cue manipulation	External judgment
		Extraversion	positive					

Appendix E. Personality Perception of Voice Parameters of the Frequency Dimension

Author	Voice Parameter	Personality	Relationship	Gender of Speaker	Type of Speech Data	Language	Study Design	Personality Assessment Method
Apple et al. (1979)	Pitch (f0 Mean)	Nervousness	positive	male	Spontaneous speech	American English	Cue synthesis	External judgment
		Persuasion	negative					
		Trustworthiness	negative					
Aronovitch (1976)	Pitch (f0 Mean)	Boldness	positive	male	Spontaneous speech	American English	Correlation study	External judgment
		Social Dominance	positive	male				
		Emotionality	positive	female				
		Extraversion	positive	male				
		Humor	positive	female				
		Kindness	positive	female				
		Maturity	negative	female				
		Self-Confidence	positive	male				
Borkowska and Paw-lowski (2011)	Pitch (f0 Mean)	Social Dominance	negative	female	Vowels	Polish	Cue synthesis	External judgment
Brown et al. (1974)	Pitch (f0 Mean)	Benevolence	negative	male	Read text	American English	Cue synthesis	External judgment
		Competence	negative					
Charoenruk and Olson (2018)	Pitch (f0 Mean)	Confidence	inverted U-shaped relationship	male	Read text	American English	Correlation study	External judgment
		Confidence	inverted U-shaped relationship	female				
		Reliability	inverted U-shaped relationship	male				
		Reliability	inverted U-shaped relationship	female				
		Trustworthiness	inverted U-shaped relationship	male				
		Trustworthiness	inverted U-shaped relationship	female				

Feinberg et al. (2005)	Pitch (f0 Mean)	Masculinity	negative	male	Vowels	American English	Cue synthesis	External judgment
Fraccaro et al. (2013)	Pitch (f0 Mean)	Dominance*	negative	male	Vowels	Canadian English	Cue manipulation	External judgment
		Dominance*	negative	female				
Guyer et al. (2019)	Pitch (f0 Mean)	Confidence	negative	male	Read text	Canadian English	Cue synthesis	External judgment
Jones et al. (2010)	Pitch (f0 Mean)	Dominance*	negative	male	Vowels	British English	Cue synthesis	External judgment
		Dominance*	negative	female				
Klofstad et al. (2015)	Pitch (f0 Mean)	Competence	negative	male	Semi-spontaneous speech	American English	Cue synthesis	External judgment
		Competence	negative	female				
Koutsoumpis and Vries (2022)	Pitch (f0 Mean)	Agreeableness	negative	male	Read text & vowels	Dutch	Correlation study	Self-ratings
		Openness to Experience	negative	female				
Krahé and Pa-pakonstan-tinou (2020)	Pitch (f0 Mean)	Feminity	positive	female	Read text	German	Cue synthesis	External judgment
		Likability	positive					
		Masculinity	negative					
Krahé et al. (2021)	Pitch (f0 Mean)	Competence	negative	female	Read text	German	Cue synthesis	External judgment
		Feminity	positive	male				
		Feminity	positive	female				
		Likability	positive	male				
		Likability	positive	female				
		Masculinity	negative	male				
Mallory and Miller (1958)	Pitch (f0 Mean)	Social Dominance	negative	female	Spontaneous speech	American English	Correlation study	Self-ratings
		Introversion	positive					
		Submissiveness	positive					

Oksenberg et al. (1986)	Pitch (f0 Mean)	Competence	positive	female	Semi-spontaneous speech	American English	Correlation study	External judgment
Oleszkiewicz et al. (2017)	Pitch (f0 Mean)	Competence	negative	male	Vowels	Polish	Cue synthesis	External judgment
		Competence	negative	female				
		Trustworthiness	negative	male				
		Trustworthiness	negative	female				
		Warmth	positive	female				
Pisanski and Rendall (2011)	Pitch (f0 Mean)	Masculinity	negative	male	Read text	Canadian English	Cue synthesis	External judgment
		Masculinity	negative	female				
Puts et al. (2007)	Pitch (f0 Mean)	Aggressiveness/ Physical Dominance	negative	male	Spontaneous speech	American English	Cue synthesis	External judgment
Puts et al. (2006)	Pitch (f0 Mean)	Social Dominance	negative	male	Spontaneous speech	American English	Cue synthesis	External judgment
		Aggressiveness/ Physical Dominance	negative	male	Spontaneous speech	American English	Cue synthesis	External judgment
Pisanski and Rendall (2011)	Pitch (f0 Mean)	Social Attractiveness	negative	male	Spontaneous speech	American English	Cue synthesis	External judgment
		Benevolence	inverted U-shaped relationship					
Rosenberg and Hirschberg (2009)	Pitch (f0 Mean)	Charisma	positive	male	Semi-spontaneous speech	American English	Correlation study	External judgment
Scherer (1978)	Pitch (f0 Mean)	Agreeableness	negative	male	Spontaneous speech	American English	Correlation study	External judgment
		Neuroticism	positive					
Tigue et al. (2012)	Pitch (f0 Mean)	Dominance*	negative	male	Spontaneous speech	American English	Cue synthesis	External judgment
		Intelligence	negative					
		Trustworthiness	negative					
	Pitch (f0 Mean)	Social Dominance	negative	male	Read text	Scottish English	Cue synthesis	External judgment

Tsantani et al. (2016)		Trustworthiness	negative	male				
		Trustworthiness	negative	female				
Tusing and Dillard (2000)	Pitch (f0 Mean)	Social Dominance	positive	male	Semi-spontaneous speech	American English	Correlation study	External judgment
van Bezooi-jen (1995)	Pitch (f0 Mean)	Independence	negative	female	Read text	Dutch	Cue synthesis	External judgment
		Independence	negative	male				
		Modesty	positive	female				
		Modesty	positive	male				
		Independence	negative	female		Japanese		
		Independence	negative	male				
		Modesty	positive	female				
		Modesty	positive	male				
Zoghaib (2017)	Pitch (f0 Mean)	Maturity	positive	female	Semi-spontaneous speech	French	Correlation study	External judgment
		Maturity	positive	male				
		Warmth	positive	female				
		Warmth	positive	male				
		Young	negative	female				
		Young	negative	male				
Addington (1968)	Pitch variability (f0 SD)	Activity/ Energy	positive	male	Read text	American English	Cue manipulation	External judgment
		Artistic	positive	male				
		Emotionality	positive	male				
		Enthusiasm	positive	male				
		Enthusiasm	positive	female				

		Feminity	positive	male				
		Humor	positive	male				
		Kindness	positive	male				
		Modesty	negative	female				
		Politeness	positive	male				
		Sensitivity	positive	male				
		Talkativeness	positive	male				
		Talkativeness	positive	female				
		Willingness to Cooperate	positive	male				
Brown et al. (1972)	Pitch variability (f0 SD)	Competence	positive	male	Read text	American English	Cue synthesis	External judgment
Brown et al. (1973)	Pitch variability (f0 SD)	Benevolence	positive	male	Read text	American English	Cue synthesis	External judgment
Brown et al. (1974)	Pitch variability (f0 SD)	Benevolence	positive	male	Read text	American English	Cue synthesis	External judgment
		Competence	positive					
Charoenruk and Olson (2018)	Pitch variability (f0 SD)	Confidence	inverted U-shaped relationship	female	Read text	American English	Correlation study	External judgment
		Confidence	negative	male				
		Reliability	negative	male				
		Reliability	negative	female				
		Trustworthiness	negative	male				
		Trustworthiness	inverted U-shaped relationship	female				
Oksenberg et al. (1986)	Pitch variability (f0 SD)	Competence	positive	female	Semi-spontaneous speech	American English	Correlation study	External judgment
Ray et al. (1991)	Pitch variability (f0 SD)	Social Attractiveness	positive	female	Read text	American English	Cue manipulation	External judgment
		Dynamic	positive					

Ray (1986)	Pitch variability (f0 SD)	Benevolence	positive	male	Read text	American English	Cue manipulation	External judgment
		Competence	positive					
		Social Dominance	negative					

Notes: *These studies do not define the dominance measured, so it is unclear whether physical or social dominance was measured. Tigue et al. (2012) measured dominance by asking participants to "choose the voice that […] sounds more dominant" (p. 211). Also, participants of Fraccaro et al. (2013) and Jones et al. (2010) compared two voices and were asked to rate which voice sounded more dominant. Levitt and Lucas (2018) measured dominance by asking participants to "rate [the voice] on a scale of 1-to-9 for dominance" (p. 402). Physical dominance can be defined as "fighting ability" (Puts et al., 2007, p. 340) or "aggression" (Puts et al., 2006, p. 284); social dominance can be described through attributes like leadership, persuasion, and "having power and influence over others" (Tsantani et al., 2016, p. 950) and is the opposite of submissiveness (Mallory & Miller, 1958; Puts et al., 2006).

Appendix F. Personality Perception of Voice Parameters of the Amplitude Dimension

Author	Voice Parameter	Personality	Relationship	Gender of Speaker	Type of Speech Data	Language	Study Design	Personality Assessment Method
Aronovitch (1976)	Loudness (Intensity level)	Activity/ Energy	positive	female	Spontaneous speech	American English	Correlation study	External judgment
		Boldness	positive	female				
		Boldness	positive	male				
		Social Dominance	positive	female				
		Social Dominance	positive	male				
		Extraversion	positive	female				
		Extraversion	positive	male				
		Kindness	positive	male				
		Maturity	positive	female				
		Self-Confidence	positive	female				
Kimble and Seidel (1991)	Loudness (Intensity level)	Assertiveness	positive	male	Spontaneous speech	American English	Correlation study	Self-ratings
		Assertiveness	positive	female				
Mallory and Mil-ler (1958)	Loudness (Intensity level)	Social Dominance	positive	female	Spontaneous speech	American English	Correlation study	Self-ratings
		Introversion	negative	female				
		Submissiveness	negative	female				
Oksenberg et al. (1986)	Loudness (Intensity level)	Competence	positive	female	Semi-spontaneous speech	American English	Correlation study	External judgment
Peng et al. (1993)	Loudness (Intensity level)	Competence	positive	male	Read text	Korean	Correlation study	External judgment
Ray (1986)	Loudness (Intensity level)	Benevolence	negative	male	Read text	American English	Cue manipulation	External judgment
		Competence	positive					
Rosenberg and Hirschberg (2009)	Loudness (Intensity level)	Charisma	positive	male	Semi-spontaneous speech	American English	Correlation study	External judgment
Scherer (1978)	Loudness (Intensity level)	Extraversion	positive	male	Spontaneous speech	American English	Correlation study	External judgment

		Neuroticism	negative					
Tusing and Dillard (2000)	Loudness (Intensity level)	Social Dominance	positive	male	Semi-spontaneous speech	American English	Correlation study	External judgment
		Social Dominance	positive	female				
Koutsoumpis and Vries (2022)	Loudness variability (Intensity range)	Conscientiousness	negative	male	Read text & Vowels	Dutch	Correlation study	Self-ratings
		Conscientiousness	negative	female				
		Openness to Experience	positive	male				
		Openness to Experience	positive	female				
Aronovitch (1976)	Loudness variability (Intensity standard deviation)	Self-Confidence	positive	male	Spontaneous speech	American English	Correlation study	External judgment
Scherer (1978)	Loudness variability (Intensity range)	Extraversion	positive	male	Spontaneous speech	American English	Correlation study	External judgment
		Neuroticism	negative					
Tusing and Dillard (2000)	Loudness variability (Intensity standard deviation)	Social Dominance	positive	male	Semi-spontaneous speech	American English	Correlation study	External judgment
		Social Dominance	positive	female				

Appendix G. Personality Perception of Voice Parameters of the Voice Quality Dimension

Author	Voice Parameter	Personality	Relationship	Gender of Speaker	Type of Speech Data	Language	Study Design	Personality Assessment Method
Addington (1968)	Breathiness	Artistic	positive	male	Read text	American English	Cue manipulation	External judgment
		Emotionality	positive	female				
		Enthusiasm	positive	female				
		Feminity	positive	female				
		Humor	positive	female				
		Idealistic	positive	female				
		Kindness	positive	female				
		Maturity	negative	female				
		Politeness	positive	female				
		Sensitivity	positive	female				
		Sophistication	negative	female				
		Young	positive	male				
		Young	positive	female				
Levitt and Lucas (2018)	Breathiness	Sexiness	positive	female	Vowels	American English	Cue manipulation	External judgment
Moore (1939)	Breathiness	Extraversion	negative	unknown (probably male)	Unknown	American English	Correlation study	Self-ratings
		Neuroticism	positive					
Pearsell and Pape (2023)	Breathiness	Agreeableness	negative	male	Read text	Canadian English	Cue manipulation	External judgment
		Agreeableness	negative	female				
		Conscientiousness	negative	male				
		Conscientiousness	negative	female				

		Extraversion	negative	male				
		Extraversion	negative	female				
		Neuroticism	negative	female				
		Neuroticism	negative	male				
Scherer (1978)	Breathiness	Assertiveness	negative	male	Spontaneous speech	American English	Correlation study	External judgment
van Borsel et al. (2009)	Breathiness	Feminity	positive	female	Vowels	Dutch	Cue manipulation	External judgment
Waaramaa et al. (2021)	Breathiness	Artistic	positive	female	Read text	Finnish	Cue manipulation	External judgment
		Artistic	positive	male				
		Determination	negative	male				
		Determination	negative	female				
		Emotionality	positive	male				
		Emotionality	positive	female				
		Persuasion	positive	male				
		Persuasion	positive	female				
		Submissiveness	positive	male				
		Submissiveness	positive	female				
		Timidity	positive	male				
		Timidity	positive	female				
		Trustworthiness	negative	male				
		Trustworthiness	negative	female				
Anderson et al. (2014)	Creakiness	Competence	negative	female	Semi-spontaneous speech	American English	Cue manipulation	External judgment
		Intelligence	negative					

		Trustworthiness	negative					
Levitt and Lucas (2018)	Creakiness	Sexiness	negative	female	Vowels	American English	Cue manipulation	External judgment
		Youthfulness	negative					
Pearsell and Pape (2023)	Creakiness	Agreeableness	negative	male	Read text	Canadian English	Cue manipulation	External judgment
		Agreeableness	negative	female				
		Conscientiousness	negative	male				
		Conscientiousness	negative	female				
		Extraversion	negative	male				
		Extraversion	negative	female				
		Neuroticism	negative	male				
		Neuroticism	negative	female				
Waaramaa et al. (2021)	Creakiness	Activity/ Energy	negative	male	Read text	Finnish	Cue manipulation	External judgment
		Activity/ Energy	negative	female				
		Enthusiasm	negative	male				
		Enthusiasm	negative	female				
		Extraversion	negative	male				
		Friendliness	negative	male				
		Friendliness	negative	female				
		Introversion	positive	female				
		Persuasion	positive	male				
		Persuasion	positive	female				
Yuasa (2010)	Creakiness	Aggressiveness/ Physical Dominance	negative	female	Spontaneous speech	American English	Cue manipulation	External judgment
		Confidence	negative					

		Hesitancy	positive					
Addington (1968)	Flatness	Activity/ Energy	negative	male	Read text	American English	Cue manipulation	External judgment
		Activity/ Energy	negative	female				
		Emotionality	negative	male				
		Emotionality	negative	female				
		Enthusiasm	negative	female				
		Enthusiasm	negative	male				
		Feminity	negative	female				
		Humor	negative	male				
		Kindness	negative	male				
		Masculinity	positive	male				
		Sensitivity	negative	female				
		Sensitivity	negative	male				
		Talkativeness	negative	male				
Moore (1939)	Harshness	Social Dominance	positive	unknown (probably male)	Unknown	American English	Correlation study	Self-ratings
		Neuroticism	negative					
Scherer (1978)	Harshness	Conscientiousness	negative	male	Spontaneous speech	American English	Correlation study	External judgment
Addington (1968)	Nasality	Activity/ Energy	negative	male	Read text	American English	Cue manipulation	External judgment
		Artistic	negative	female				
		Artistic	negative	male				
		Carefulness	negative	male				
		Carefulness	negative	female				

		Emotionality	negative	female				
		Enthusiasm	negative	male				
		Enthusiasm	negative	female				
		Intelligence	negative	male				
		Intelligence	negative	female				
		Maturity	negative	male				
		Modesty	positive	male				
		Modesty	positive	female				
		Neuroticism	positive	male				
		Politeness	negative	male				
		Politeness	negative	female				
		Sophistication	negative	male				
		Sophistication	negative	female				
Moore (1939)	Nasality	Social Dominance	negative	unknown (probably male)	Unknown	American English	Correlation study	Self-ratings
		Neuroticism	positive					
Scherer (1978)	Nasality	Extraversion	positive	male	Spontaneous speech	American English	Correlation study	External judgment
		Neuroticism	negative					
Waaramaa et al. (2021)	Nasality	Enthusiasm	negative	male	Read text	Finnish	Cue manipulation	External judgment
		Enthusiasm	negative	female				
		Persuasion	positive	male				
		Persuasion	positive	female				
		Trustworthiness	negative	male				
		Trustworthiness	negative	female				

Addington (1968)	Resonance	Activity/ Energy	positive	male	Read text	American English	Cue manipulation	External judgment
		Artistic	positive	male				
		Artistic	positive	female				
		Emotionality	positive	female				
		Enthusiasm	positive	male				
		Enthusiasm	positive	female				
		Humor	negative	female				
		Idealistic	positive	female				
		Modesty	negative	male				
		Modesty	negative	female				
		Sophistication	positive	male				
		Talkativeness	positive	female				
Mallory and Miller (1958)	Resonance	Social Dominance	positive	female	Spontaneous speech	American English	Correlation study	Self-ratings
		Introversion	negative					
		Submissiveness	negative					
Scherer (1978)	Resonance	Conscientiousness	positive	male	Spontaneous speech	American English	Correlation study	External judgment
		Neuroticism	negative					
Feinberg et al. (2005)	Resonance (Formant dispersion)	Masculinity	negative	male	Vowels	American English	Cue synthesis	External judgment
Levitt and Lucas (2018)	Resonance (Formant dispersion)	Dominance*	negative	female	Vowels	American English	Cue manipulation	External judgment
		Sexiness	positive					
		Youthfulness	positive					
	Resonance	Masculinity	negative	male	Read text	Canadian English	Cue synthesis	External judgment

Pisanski and Rendall (2011)	(Formant dispersion)	Masculinity	negative	female				
Puts et al. (2007)	Resonance (Formant dispersion)	Aggressiveness/ Physical Dominance	negative	male	Spontaneous speech	American English	Cue synthesis	External judgment
Feinberg et al. (2005)	Resonance (Formant dispersion)	Masculinity	negative	male	Vowels	American English	Cue synthesis	External judgment
Addington (1968)	Roughness	Activity/ Energy	negative	female	Read text	American English	Cue manipulation	External judgment
		Artistic	negative	female				
		Carefulness	negative	female				
		Emotionality	negative	female				
		Enthusiasm	negative	female				
		Feminity	negative	female				
		Idealistic	negative	male				
		Intelligence	negative	female				
		Maturity	positive	male				
		Modesty	positive	female				
		Neuroticism	positive	female				
		Neuroticism	negative	male				
		Politeness	negative	female				
		Sophistication	positive	male				
		Sophistication	negative	female				
		Talkativeness	negative	female				
		Young	negative	male				
	Roughness (HNR)	Openness to Experience	negative	male	Read text & Vowels	Dutch	Correlation study	Self-ratings

Koutsoumpis and Vries (2022)		Openness to Experience	negative	female				
Wu et al. (2021)	Roughness (HNR)	Approachability	negative	male	Read text	Chinese	Correlation study	External judgment
Scherer (1978)	Sharpness	Agreeableness	negative	male	Spontaneous speech	American English	Correlation study	External judgment
		Neuroticism	positive					
Addington (1968)	Tension	Activity/ Energy	positive	female	Read text	American English	Cue manipulation	External judgment
		Artistic	negative	male				
		Carefulness	negative	female				
		Carefulness	negative	male				
		Emotionality	positive	female				
		Enthusiasm	positive	female				
		Feminity	positive	female				
		Intelligence	negative	female				
		Intelligence	negative	male				
		Kindness	negative	male				
		Maturity	negative	female				
		Neuroticism	positive	male				
		Neuroticism	positive	female				
		Politeness	negative	male				
		Sensitivity	negative	male				
		Sensitivity	positive	female				
		Sophistication	negative	female				

		Willingness to Cooperate	negative	male				
		Young	negative	male				
		Young	positive	female				
Moore (1939)	Tension	Social Dominance	positive	unknown (probably male)	Unknown	American English	Correlation study	Self-ratings
		Neuroticism	negative					
Waaramaa et al. (2021)	Tension	Determination	positive	female	Read text	Finnish	Cue manipulation	External judgment
		Determination	positive	male				
		Social Dominance	positive	female				
		Social Dominance	positive	male				
		Emotionality	negative	female				
		Emotionality	negative	male				
		Friendliness	negative	male				
		Friendliness	negative	female				
		Willingness to Cooperate	negative	female				
		Willingness to Cooperate	negative	male				
Addington (1968)	Thinness	Carefulness	negative	female	Read text	American English	Cue manipulation	External judgment
		Feminity	positive					
		Humor	positive					
		Intelligence	negative					
		Kindness	positive					
		Maturity	negative					
		Modesty	positive					
		Neuroticism	negative					

		Sensitivity	positive					
		Sophistication	negative					
		Young	positive					
Scherer (1978)	Thinness	Agreeableness	negative	male	Spontaneous speech	American English	Correlation study	External judgment
		Neuroticism	positive					

Notes: *This study does not define the dominance measured, so it is unclear whether physical or social dominance was measured. Tigue et al. (2012) measured dominance by asking participants to "choose the voice that […] sounds more dominant" (p. 211). Also, participants of Fraccaro et al. (2013) and Jones et al. (2010) compared two voices and were asked to rate which voice sounded more dominant. Levitt and Lucas (2018) measured dominance by asking participants to "rate [the voice] on a scale of 1-to-9 for dominance" (p. 402). Physical dominance can be defined as "fighting ability" (Puts et al., 2007, p. 340) or "aggression" (Puts et al., 2006, p. 284); social dominance can be described through attributes like leadership, persuasion, and "having power and influence over others" (Tsantani et al., 2016, p. 950) and is the opposite of submissiveness (Mallory & Miller, 1958; Puts et al., 2006). HNR = harmonic-to-noise ratio.

Appendix H. Personality Perception of Linear Combination of Voice Parameters

Author	Linear Combination of Voice Parameters	Personality (Linear Combination leads to the Perception of …)	Gender of Speaker	Type of Speech Data	Language	Study Design	Personality Assessment Method
Feinberg et al. (2005)	Decreasing pitch + decreasing Df	Decreased Masculinity	male	Vowels	American English	Cue synthesis	External judgment
McAleer et al. (2014)	Increasing pitch + decreasing Df	Increased Social Dominance	female	Read text	Scottish English	Correlation study	External judgment
	Increasing pitch range + decreasing HNR + decreasing glide	Increased Trustworthiness					
	Decreasing pitch + decreasing Df + decreasing HNR + decreasing spectral slope	Increased Social Dominance	male				
	Increasing pitch + decreasing HNR	Increased Trustworthiness					
Pisanski and Rendall (2011)	Decreasing pitch + decreasing Df	Increased Masculinity	female	Read text	Canadian English	Cue synthesis	External judgment
	Decreasing pitch + decreasing Df	Increased Masculinity	male				
Ray (1986)	Decreasing pitch variability + increasing speaking rate	Decreased Benevolence	male	Read text	American English	Cue manipulation	External judgment
	Increasing pitch variability + decreasing loudness	Increased Benevolence					
	Decreasing pitch variability + decreasing speaking rate	Decreased Competence					
	Increasing pitch variability + increasing speaking rate	Increased Competence					
Riding et al. (2006)	Increasing pitch variability + increasing speaking rate	Increased Benevolence	male	Spontaneous speech	American English	Cue synthesis	External judgment
Wu et al. (2021)	Increasing pitch + increasing spectral slope	Increased Capability	female	Read text	Chinese	Correlation study	External judgment
	Decreasing HNR + increasing spectral slope	Increased Capability	male				

Notes: Df = formant dispersion; HNR = harmonic-to-noise ratio.

Appendix I. Journal Overview

Journal	No. of Articles	Research Areas	Impact Factor (2022)	SJR (2023)
American Speech	1	Linguistics & Language; Communication	0.4	0.587
Animal Behaviour	4	Ecology, Evolution, Behavior & Systematics	2.5	0.924
Communication Monographs (former: Speech Monographs)	4	Linguistics & Language; Communication	2.5	1.261
Communication Quarterly	1	Communication	1.8	0.483
European Journal of Social Psychology	1	Social Psychology	3.9	1.630
Evolution and Human Behavior	3	Experimental & Cognitive Psychology	5.1	1.562
Field Methods	1	Social Sciences	1.7	0.467
Frontiers in Communication	1	Social Sciences	2.4	0.589
Human Communication Research	1	Developmental & Educational Psychology; Linguistics & Language; Communication	5.0	2.034
Journal of Cross-Cultural Psychology	2	Social Psychology; Cultural Studies	3.5	0.992
Journal of Individual Differences	1	Psychology	1.4	0.424
Journal of Language and Social Psychology	1	Social Psychology; Linguistics & Language	2.1	1.246
Journal of Nonverbal Behavior	3	Social Psychology	2.9	0.574
Journal of Personality and Social Psychology	2	Social Psychology	7.6	3.610
Journal of Product and Brand Management	1	Marketing	4.4	1.685
Journal of Speech, Language, and Hearing Research (former: Journal of Speech Disorders)	1	Linguistics & Language	2.6	0.827
Journal of Voice	2	Speech & Hearing	2.2	0.578
Language and Communication	1	Linguistics & Language; Communication; Social Psychology	1.9	0.667
Language and Speech	2	Linguistics & Language	1.8	0.625
Perception	1	Computer Science; Experi-mental & Cognitive Psychology	1.7	0.584
Personality and Social Psychology Bulletin	2	Social Psychology	4.0	2.325
PLOS ONE	3	Multidisciplinary	3.7	0.839
Pollack Periodica	1	Computer Science; Engineering	0.8	0.288
Psychology of Language and Communication	1	Linguistics & Language; Communication; Experimental & Cognitive Psychology	0.4	0.241
Psychonomic Bulletin and Review	1	Developmental & Educational Psychology; Experimental & Cognitive Psychology	4.7	1.753
Public Opinion Quarterly	1	Social Sciences	3.4	1.636
Sex Roles	1	Developmental & Educational Psychology; Gender Studies	4.6	1.216
Social Psychology	1	Social Psychology	1.8	0.668
Speech Communication	1	Computer Science; Linguistics & Language	3.2	0.769
The Journal of Social Psychology	2	Social Psychology	2.1	0.750
The Journal of the Acoustical Society of America	4	Acoustics & Ultrasonics	2.4	0.687

Notes: SJR = SCImago Journal Rank.

Essay 2

Can You Hear My Personality?

A Conceptualization of a Brand Voice Based on Brand Personality

Author: Olga Bosak

Abstract

The increasing prevalence of voice-based technology, such as voice assistants, has enabled more natural, frequent, and integrated brand communication with customers. Brands enter into a dialogue with their customers, thereby creating a brand experience that resembles human interaction. These interactions are usually limited to purely auditory communication, as most voice-based devices lack visual content. Consequently, the voice plays a significant role, as the brand's personality is perceived exclusively through the voice. However, how a voice must sound so listeners perceive a desired brand personality remains unclear. The present study addresses this gap by being the first empirical study to examine the perception of brand personalities through voice. Based on an exploratory approach, this study shows that listeners can perceive the brand personality dimensions of sincerity, sensitivity, and excitement through voice alone. Further, the developed brand voice personality model (BVP-Model) identifies the key objectively measurable voice parameters that help to design a female and male voice to perceive these brand voice personality dimensions. The results guide marketers and voice technology specialists in translating a brand personality into a voice and significantly contribute to vocal personality perception and voice marketing research.

Keywords: brand personality, brand voice, personality perception, voice user interfaces, voice assistants, voice AI, voice marketing

1. Introduction

In the past, consumers have typically interacted with technological devices by physically clicking, typing, or swiping (Packard & Berger, 2024). Now, consumers use voice-based interactions for tasks such as searching for information, shopping, obtaining directions, or making reservations (Melzner, Bonezzi, and Meyvis, 2022). Voice-based conversational agents and voice assistants, such as Amazon's Alexa, Google Assistant, and Apple's Siri, facilitate a user experience that closely resembles human interaction due to their socially adapted behavior, personality, and capacity for independent and interactive communication (Wagner & Schramm-Klein, 2019). From a corporate standpoint, integrating voice marketing into the overall marketing and branding strategy is highly advisable, especially given the widespread adoption of voice technologies in households worldwide. The substantial increase in global revenue of smart speakers with voice assistants installed from \$25.2 billion in 2018 to \$43.7 billion in 2023, a 73.4% rise within five years, evidences the rapid diffusion of voice technology (Statista Market Insights, 2023).

The emergence of voice artificial intelligence (AI)[1] presents a new opportunity for marketers to shape their brand image and personality, and thus to strengthen customer relationships and build brand equity (Guha et al., 2023; H. Lee & Cho, 2020; Zoghaib, 2017). The largest providers of smart speakers and voice AI, Amazon and Google, offer indirect and direct ways to advertise on their devices. Users of smart speakers can listen to radio, songs, or podcasts on audio streaming providers, which include advertising. Additionally, Alexa users can directly receive personalized product recommendations based on their previous shopping behavior (Hardesty, 2019). Further, companies can develop voice applications to provide users with entertainment, information, or services (Amazon, 2022). Brands thus enter into a dialogue with their customers,

[1] voice/ voice-based AI = AI technologies enabled to perform speech recognition and natural language processing (NLP) to engage in natural dialogues with users (Wang et al., 2023)

creating a new brand experience and forming new brand associations (Hörner, 2023). The growing prevalence of voice AI in everyday life has not only facilitated bidirectional and interactive communication between consumers and brands but has also made it a more natural, frequent, and integrated aspect of everyday routines (Jurafsky & Martin, 2020; V. Kumar et al., 2016).

While numerous companies want to engage in voice marketing, determining the most appropriate voice for their brand is a major initial challenge. With voice AI, user and system interaction is usually limited to purely auditory communication (Packard & Berger, 2024). Research in phonetics demonstrates the importance of voice and specific voice parameters, such as pitch, loudness, and intonation, in shaping speakers' perception (Schweinberger & Zäske, 2019). In addition to the verbal content, a speaker conveys non-verbal content to the listener, such as the emotional and motivational state (Gobl & Chasaide, 2003), physiological cues (Krauss, Freyberg, and Morsella, 2002), and their identity and personality (Zäske et al., 2020). For instance, a person with a low-pitched voice is likely to be perceived as more competent, confident, and trustworthy (Oleszkiewicz et al., 2017; Rodero, 2013); a person with a relatively fast speaking rate is likely to be perceived as more extroverted and ambitious (Addington, 1968; 1985). Consequently, the voice of a spokesperson, in addition to the content provided, will also convey the brand's personality in voice-based AI interactions (Klasmeyer & Sendlmeier, 1997; Zäske et al., 2020). Therefore, marketers must identify a voice that fits their brand so that customers can also perceive the intended brand personality in auditory communication. Despite the importance of selecting an appropriate brand voice, marketers typically rely on their intuition and experience for brand voice selection (Dahl, 2010; Wiener & Chartrand, 2014). Although there is a substantial corpus of empirical evidence on the perception of human personalities through voice, there is no systematic approach to identifying the appropriate voice for conveying the personality of a brand.

Building on the existing body of perceptual research, the present study addresses this gap by suggesting a brand voice personality model (BVP-Model) that determines how brand personalities are perceived through voice. Therefore, this study's *initial research question* concerns identifying which brand personalities can be perceived through voice alone, i.e., brand voice personalities, given that previous perceptual studies have primarily focused on human personalities. In doing so, a brand voice personality scale (BVP-Scale) is developed to measure the brand personalities that are perceived by voice alone. In light of these findings, the *second research question* concerns determining which combinations of voice parameters in female and male voices induce the perception of the identified brand voice personalities to form the final model. To this end, a comprehensive correlational study is conducted in which naïve listeners evaluate a substantial corpus of voices regarding perceived brand personalities. This study's results provide marketers, sound designers, and voice engineers with a systematic approach for selecting or designing a brand voice aligned with a desired brand personality.

Further, this study makes three contributions to vocal personality perception and voice marketing research. Primarily, it is the first empirical study to examine the perception of brand personalities through voice compared to previous research focusing on vocal human personality perception. Secondly, this study examines the influence of linear combinations of voice parameters on brand voice personality perceptions, thereby reacting to the call for multivariate analyses of voice parameters in perceptual studies (Feinberg et al., 2005; McAleer et al., 2014; Pisanski & Rendall, 2011; Wu et al., 2021). Finally, this study's exploratory approach considers only objectively measurable voice parameters and their acoustical measures, ensuring the comparability of results and expanding the knowledge on voice perceptions.

The remainder of this article is structured as follows: First, related work on vocal personality perception is reviewed, and the scope of the study is defined. This review forms the basis for developing the conceptual framework of the BVP-Model. Subsequently, the procedure for determining which brand personality traits are associated with voices is outlined. Afterward, the results of the development of the BVP-Scale are presented and interpreted. Then, the results of the BVP-Model development and their interpretation are presented. Finally, the study concludes with the managerial and theoretical implications and recommendations for future research based on the limitations of this study.

2. Background and Conceptual Development

Researchers have conducted studies on vocal personality perception for almost a century, (Allport & Cantril, 1934; Moore, 1939; Pear, 1931). With the development of advanced technologies and software for speech processing analysis, next to progress in phonetics and psychology, research on voice perception has gained significant attention from the 1960s to the 1980s. This extensive research on vocal personality perception shows that the paraverbal content of speech significantly influences the perception of personality traits, such as competence, dominance, and extraversion (Kreiman & Sidtis, 2011; Mehrabian & Ferris, 1967). Paraverbal content refers to the acoustic parameters of the voice surrounding the semantic content of a message, such as speech rate, pitch, or intensity, which are expressed differently in each voice (Ketrow, 1990).

All humans share a comparable structure of the vocal tract system. Yet, differences in the length of vocal folds and slight variations in the acoustic parameters around a mean contribute to the unique characteristics of each voice. These variations leave every person with an individual vocal signature (Belin et al., 2011; Rodero, 2013). Despite the uniqueness of voices, phoneticians and psychologists have discovered correlations between several voice parameters and specific personality traits. Researchers suggest that voice stereotypes exist, meaning that single vocal parameters or their combination cause listeners to perceive the same personality in different voices (for a comprehensive overview, please refer to the systematic literature review on vocal personality perception in **Essay 1**; Addington, 1968; Kramer, 1964).

2.1. Related Work and Scope of Study

While numerous studies on the perception of personality through voice exist, they differ in various aspects, making them comparable to a limited extent. The key differences relate to the *study design* and *assessment method* used to investigate perceived personalities. Phonetic experiments can be designed as correlation, cue manipulation, or cue synthesis studies (Brown et al., 1985; Riding et al., 2006). In correlational study designs, voice parameters are measured without manipulation. In cue manipulation studies, speakers subjectively manipulate their voices, e.g., the speaker raises their pitch subjectively. This approach implies that speakers may also simultaneously alter other vocal parameters. In cue synthesis studies, voices are manipulated through software, e.g., the speaker's mean pitch is raised by 20%. Researchers achieve greater experimental control with cue synthesis because individual voice parameters can be manipulated while others can be held constant (Brown et al., 1985; Riding et al., 2006). Additionally, two different assessment methods of perceived personalities exist: external judgments of (naïve) listeners or self-ratings of the speaker. Research suggests that the accuracy and validity of personality perception may vary depending on the assessment method used. Self-ratings of personality traits may yield different results than external judgments due to values, experiences, and self-perceptions that are not apparent to external individuals (Koutsombogera et al., 2020).

To determine how brand personalities are perceived through voice, this study employs a correlational design and uses external judgments of personalities from naïve listeners for two reasons. Firstly, the aim is to investigate the external perception of brand personalities from a consumer's point of view. Thus, self-perceptions are not relevant. Secondly, a substantial number of voices is examined to assess which individual voice parameters are crucial for perceiving specific brand personalities. Considering these methodological decisions, a summary of related correlation studies employing external assessment methods on the perception of voice personality is presented below to derive the scope of this study (see **Table 1**).

Table 1. Summary of Related Work on Vocal Personality Perception

Study	A. Language	B. Analysis of Voice Parameters		C. Measurement of Voice Parameters		D. Gender of Speaker		E. Investigated Personality	
		univariate	**multivariate**	**subjectively**	**objectively**	**female**	**male**	**human**	**brand**
Aronovitch (1976)	American English	x			x	x	x	x	
Scherer (1978)	American English	x		x			x	x	
Oksenberg et al. (1986)	American English	x		x		x		x	
Peng et al. (1993)	American English & Korean	x		x			x	x	
Tusing and Dillard (2000)	American English	x			x	x	x	x	
Rosenberg and Hirschberg (2009)	American English	x			x		x	x	
McAleer et al. (2014)	Scottish English		x		x	x	x	x	
Zoghaib (2017)	French	x			x	x	x	x	
Charoenruk and Olson (2018)	American English	x			x	x	x	x	
Wu et al. (2021)	Chinese	x			x	x	x	x	
This Study	**German**		**x**		**x**	**x**	**x**		**x**

Notes: This summary of related work shows only correlational studies of vocal personality perception in which external judgments of naïve listeners assessed personalities.

Research on vocal personality perception has mainly been conducted in American English. One possible reason for the underrepresentation of other languages could be the general predominance of the (American) English language and the WEIRD countries (western, educated, industrialized, rich, and democratic) in scientific publications (Atari & Henrich, 2023; Blasi et al., 2022; Bryant, 2022). The perception of personality traits in voices differs according to cultural context though, with listeners from different cultural groups attributing distinctive personality characteristics to the same speaker (Kreiman & Sidtis, 2011). The reason for this is that the cultural patterns and social interaction rules that are learned and followed in one culture are not necessarily the same as those learned and followed in another culture (Brown et al., 1975; Hogan & Bond, 2009). Therefore, the insights on vocal perception of personality are not transferable from one culture to another, and research must be conducted separately in each culture or subculture. To contribute to the research on languages other than American English, this study will first focus on German voices with the goal of replicating the study in different countries in the future (see **Table 1A**).

Studies analyzing the effects of voice on personality perception have generally focused on identifying relationships between individual voice parameters and personalities, i.e., univariate analysis of voice parameters. They primarily did not consider interaction effects between multiple voice parameters, i.e., no multivariate analysis was done. Among related studies, only McAleer et al. (2014) conducted a multivariate analysis as they investigated linear combinations of voice parameters on personality judgments. The absence of sophisticated speech recognition and processing software may have constrained researchers from examining the interaction effects of acoustic parameters. Nevertheless, given that voice parameters covary in natural speech, it is crucial to investigate which combination of parameters induces which personality perceptions (Apple et al., 1979). Thus, this study examines the interaction effects of voice parameters and their relationship to personality judgments, thereby addressing the identified need for multivariate analyses in existing perceptual research (see **Table 1B**; Addington, 1968; Apple et al., 1979; Gobl & Chasaide, 2003).

Most studies on associations of voice qualities, such as breathiness or harshness, with personality traits rely on ratings from experienced or inexperienced judges. The voice parameters are thus not measured objectively but rely on subjective ratings, making it difficult to generalize the results. For instance, Scherer (1978) examines six voice qualities that are evaluated through perceptual judgments of females who were not professional phoneticians. In the related work, only McAleer et al. (2014) and Wu et al. (2021) examine objectively measurable voice qualities like glide, spectral slope, or harmonics-to-noise ratio (HNR). Thus, to provide a comprehensive picture of a perceived personality's voice profile and make results generalizable and comparable, the present study only considers objectively measurable voice parameters, including acoustical measures of voice qualities (see **Table 1C)**.

Research on vocal personality perception shows that personalities can be perceived differently based on the speaker's gender (Trouvain et al., 2021). These differences may be due to gender stereotypes that arise through socialization within a culture or ethnicity. In a cue manipulation study, Addington (1968) finds that male personalities were perceived in terms of physical and emotional power, while female personalities were perceived in terms of social skills. Among related studies, Aronovitch (1976) observes that a fast speaking rate in females increased perceptions of boldness, whereas a fast speaking rate in males led to decreased perceptions of boldness. Also, McAleer et al. (2014) find gender-related personality perceptions, as the voice profiles for the perception of dominance and trustworthiness differed between female and male voices. Consequently, as the perception of personality through voice depends on the speaker's gender, this study considers female and male voices separately to test for gender-specific effects on brand personality perception (see **Table 1D)**.

Finally, the review of related work reveals that no related study deals with brand personalities but human personalities.[2] Even though researchers have developed brand personality traits through the anthropomorphization of brands—that is, ascribing human characteristics to a brand—some differences remain between brands and humans (Azoulay & Kapferer, 2003; Caprara et al., 2001; A. Kumar, 2018). Considering these differences might influence the design of a brand's voice. In developing the first brand personality scale (BPS), Aaker (1997) shows that brand personality dimensions exist that humans desire yet do not possess. Of her scale's five brand personality dimensions, the two dimensions, sophistication and ruggedness, showed no correlations with one of the Big Five human personalities (openness to experience, consciousness, extraversion, agreeableness, neuroticism). The assumption is that brands with these personalities would exaggerate specific characteristics, such as sexiness and glamour for prototypical sophisticated brands or strength and masculinity for rugged brands (Aaker, 1997). However, such extremes appeared to be the exception when considering human personalities. Consequently, this study is the first to employ brand personality traits in the context of vocal personality perception (see **Table 1E)**.

2.2. Conceptual Framework

Building on previous research into the perception of voice personality in humans, this study investigates the application of personality perception to brand voices. The aim is to conceptualize a model that indicates the precise combination of voice parameters that will result in the perception of a specific brand personality. This brand voice personality model (BVP-Model) shall assist marketers and speech technology specialists in translating a brand personality into a voice.

[2] The lack of research on the vocal perception of brand personalities is evident in the related work of this study and the systematic literature review presented in Essay 1. Consequently, despite nearly 85 years of research on the vocal perception of personality, only human personalities have been examined.

Since this study is the first to examine the perception of brand personalities through voice, it needs first to determine whether and which existing brand personalities can be perceived through voice alone. Therefore, the first research question (RQ) to be answered is:

RQ1: Which brand personalities can be perceived through voice?

Accordingly, this study examines several brand personality conceptualizations derived from the generalizable BPSs that can be applied to various product categories, countries, and cultures. These scales include (1) Aaker's (1997) BPS, which has gained extensive use and acceptance in marketing research due to its high reliability and validity (Aaker et al., 2001). Moreover, (2) the BPS of Geuens et al. (2009) is utilized due to its cross-cultural applicability, as this scale was already successfully tested in ten countries, including Germany. Additionally, (3) the BPS of Grohman (2009) complements the two scales above, as this scale measures gender dimensions of brand personality and, therefore, encompasses gender-specific traits. As a first step, the question of whether the established three BPSs can be replicated in a study on brand perceptions transported through voice is tested. If not, a scale for measuring brand personalities perceived through voice alone is necessary. In this study, the term "brand voice personalities" is used to describe brand personalities perceived through voice alone. This is done to differentiate them from those brand personalities that have been developed in BPSs and are typically based on visual or multimodal cues.

The next step is investigating how a voice must sound for a specific brand voice personality to be perceived. To this end, the identified brand voice personalities of the first research question are examined in relation to linear combinations of objectively measurable voice parameters. The voice parameters of interest in this study are derived from previous findings in psychology and phonetics (see following subsection). Considering the potential impact of the speaker's gender on personality judgments, a distinction is made between female and male voices in investigating brand voice personality perceptions. Accordingly, the second research question is:

RQ2: Which voice parameters in females and males induce perceptions of brand voice personalities?

The results of the second research question form the final BVP-Model. **Figure 1** presents this study's conceptual framework.

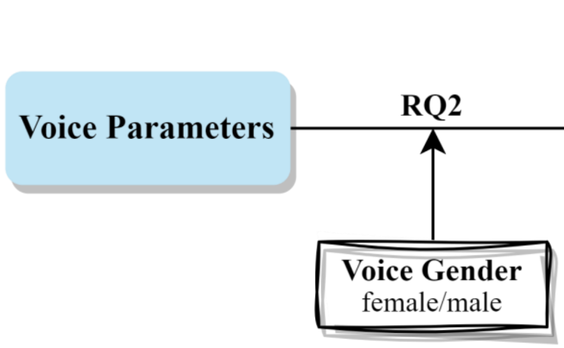


Figure 1. Conceptual Framework of Brand Voice Personality Model (BVP-Model)
Notes: RQ1: Which brand personalities can be perceived through voice? **RQ2:** Which voice parameters in females and males induce perceptions of brand voice personalities? RQ = research question.

2.2.1. Voice Parameters for Personality Perception

Regarding the second research question, 15 voice parameters and their acoustical measures[3] are identified to investigate which parameters are relevant for brand personality perceptions (see **Table 2**). All parameters are chosen to reflect different aspects of the production and perception of the voice and are frequently investigated in previous perceptual studies. These parameters can be described according to the four distinct soundwave dimensions: *timing*, *amplitude*, *frequency*, and spectral features of a voice, which will also be referred to as the *voice quality* dimension in the remainder of this article (Hildebrand et al., 2020; Jurafsky & Martin, 2020). Soundwaves are sound signals produced through the vocal fold vibrations when someone is speaking (Hildebrand et al., 2020).

[3] The term "acoustical measure" refers to the physical aspects of sound signal. In contrast, the term "voice parameter" describes the subjective qualities of a sound as perceived by listeners. Consequently, voice parameters represent a qualitative concept instead of a quantitative one (Hildebrand et al., 2020).

Table 2. Voice Parameters for Personality Perception

Soundwave Dimension	Acoustical Measure (Metric)	Listener's Perception
Timing	Speaking Rate (syl/s) Average Silent Pause Duration (s) Silent Pause Frequency per Minute (n/m)	Fluency of Speech
	Articulation Rate (syl/s)	Velocity of Speech
Amplitude	Intensity Variability (SD; dB)	Loudness Variability (Intonation)
Frequency	Fundamental Frequency Mean (f0; Hz)	Pitch
	Fundamental Frequency Standard Deviation (SD f0; Hz) Fundamental Frequency Range (f0-max – f0-min; st)	Pitch Variability (Intonation)
Voice Quality	h1-h2 (dB)	Creakiness Breathiness
	Spectral Slope Spectral Tilt	Creakiness Breathiness Loudness (Brightness)
	Smoothed Cepstral Peak Prominence (CPPS; dB)	Breathiness
	Harmonics-to-Noise Ratio (HNR; dB) Jitter (%) Shimmer (%)	Roughness/ Hoarseness

Notes: syl = syllable; s = second; m = minute; n = number; SD = standard deviation; Hz = Hertz; st = semitones; dB = Decibel.

To better understand the voice dimensions and parameters, please consider the visual representation of the soundwaves of utterances of a female and male speaker in **Figure 2**. Both persons speak the sentence “no answer is an answer, too “ in German, consisting of six words and ten syllables. An oscillogram illustrates the amplitude of an acoustic signal over time, in which amplitude values are normalized to -1.0 (representing the maximum negative sound energy) and 1.0 (representing the maximum positive sound energy). A spectrogram depicts the signal’s frequency or a combination of frequency and amplitude over time. (Hildebrand et al., 2020)

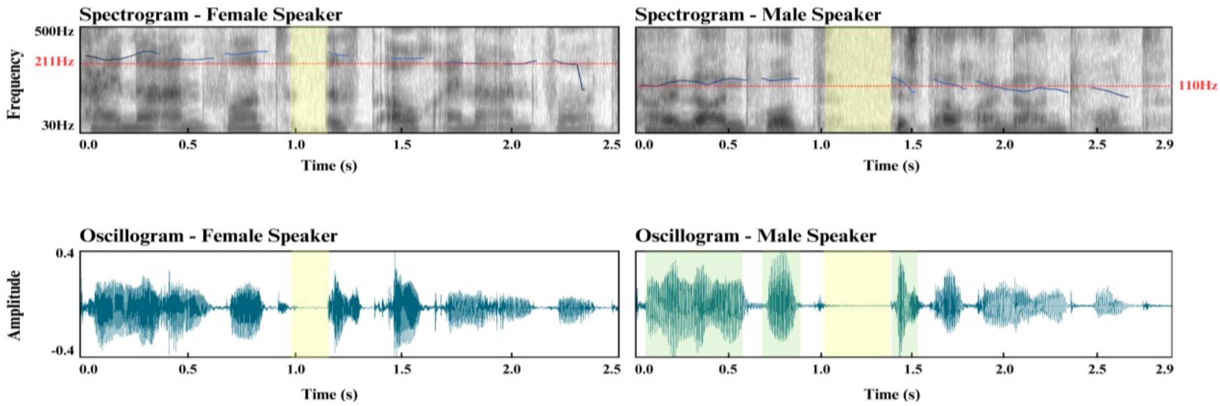


Figure 2. Spectrogram and Oscillogram of Female and Male Utterances
Notes: The utterance is the German translation of "no answer is an answer, too. " An unfilled pause is highlighted in yellow. Greater amplitude values are highlighted in green. The mean pitch is indicated with a dotted red line. Both voices were taken from the Jena Speaker Set (JESS): "Skeine_fAn25" for the female speaker; "Skeine_mHD68" for the male speaker (see Appendix A; Zäske et al., 2019).

The **timing** dimension is the duration or length of a soundwave (Hildebrand et al., 2020). *Speaking rate* refers to the number of words or syllables spoken per unit of time, including voice breaks (Peterson et al., 1995; Tusing & Dillard, 2000). *Articulation rate*, on the other hand, refers to the number of words or syllables spoken per unit of time, excluding voice breaks (Street & Brady, 1982). Voice breaks are filled or unfilled pauses made during a speech. Unfilled pauses are also called silent pauses, as they accompany a quiet inhalation, exhalation, or swallowing (Conrad et al., 2008; Koutsoumpis & Vries, 2022; Rodero, 2012). The acoustical measure of *silent pause duration* describes the duration of a silent pause, while the *frequency of silent pauses* describes how often pauses occur per unit of time. Acoustical measures such as speaking rate, silent pause duration, and frequency influence the perception of speech fluency. The articulation rate influences the perception of speech velocity, as it does not include any voice breaks. In **Figure 2**, the female speaks faster due to her articulation rate of 4.2syl/s; the articulation rate of the male speaker is 3.8syl/s. Further, the female speaker makes a shorter silent pause (.1s) than the male speaker (.3s; highlighted in yellow). Thus, the female speech is more fluent (speaking rate: 4syl/s) than the male speech (speaking rate: 3.4syl/s).

Sound intensity is measured in decibels (dB) and represents the **amplitude** of a soundwave per unit area (Hildebrand et al., 2020; Hodges-Simeon et al., 2010; Tusing & Dillard, 2000). Loudness is defined as the perceived amplitude of a voice. As absolute intensity levels and intensity means are relatively static, it is more meaningful to consider the *standard deviation of individual intensity levels* in phonetic analyses. Listeners perceive this as loudness variability, which can indicate intonation describing how monotone or dynamic a voice is perceived (Scherer, 1974). In **Figure 2**, both speakers exhibit similar loudness variability, as both amplitudes are situated between the range of -.4 and .4. However, the male speaker demonstrates slightly greater loudness variability, as his amplitude values are higher at the beginning and also around .8s and 1.5s into his speech (highlighted in green). In direct comparison, the male utterance would be perceived as more dynamic regarding loudness.

Frequency refers to a soundwave's frequency, measured in Hertz (Hz). The fundamental frequency (f0) describes the number of vibrations per second that the vocal folds make to produce a vocalization (Hildebrand et al., 2020; Koutsoumpis & Vries, 2022). Pitch is the perceptual representation of a speaker's fundamental frequency (f0). Acoustical measures of interest in this dimension include the *mean of f0, the standard deviation of f0, and the range of f0,* which is the difference between the highest (f0-max) and the lowest pitch level (f0-min; Rodero, 2017). The last two acoustical measures refer to the rise and fall of f0 over an utterance. They are perceived as pitch variability, which, like loudness variability, indicates intonation. As males are usually larger in body size than females, their vocal folds are typically larger, resulting in lower f0 values (Frühholz & Belin, 2019). This fact can be observed in **Figure 2**, as the male speaker's mean pitch is 110Hz and the female speaker's is 211Hz (indicated with a dotted red line). The female speaker exhibits a greater pitch variability (f0 SD = 27.3Hz), primarily due to her pitch lowering at the end of the sentence to 95Hz. In contrast, the male speaker shows a relatively monotone variability in his pitch with a f0 SD of 13.5Hz. In direct comparison, the female utterance would be perceived as more dynamic in pitch.

The **voice quality** dimension evaluates the spectral features of a soundwave and measures the level of periodicity and the degree of vocal instability during speech production (Hildebrand et al., 2020). The voice quality acoustical measures of interest in this study primarily indicate perceived breathiness, creakiness, or roughness/ hoarseness. These phonation types are produced through irregular vocal fold vibrations, typically absent in modal voice (Klasmeyer & Sendlmeier, 1997). Modal voice is defined as having regular and synchronized vocal fold vibrations and moderate laryngeal tension, which is also referred to as "chest voice" (see **Figure 3A**; Clark et al., 2007; Pompino-Marschall, 2009).

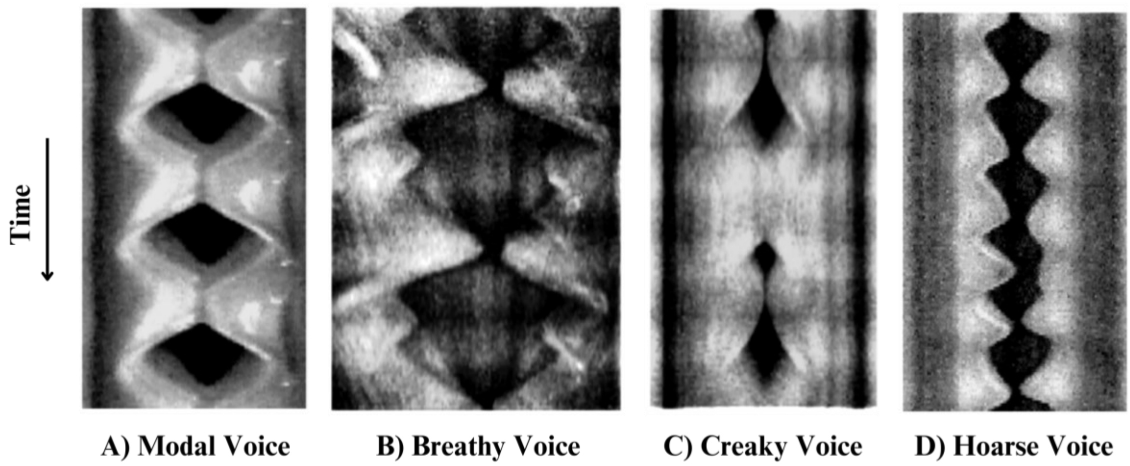


Figure 3. Vibratory Vocal Folds Pattern of a: A) Modal; B) Breathy; C) Creaky; and D) Hoarse Voice
Source: Figures adapted from Svec et al. (1999, pp. 92–93).

Breathiness in voice occurs when the vocal folds do not close completely, resulting in audible expiration noises (see **Figure 3B**; Gobl & Chasaide, 2003; Klasmeyer & Sendlmeier, 1997). Creakiness in voice is produced through a very short time opening of the vocal folds and abrupt irregular vibrations of the vocal folds, resulting in audible cracking and popping noises (see **Figure 3C**; Clark et al., 2007; Klasmeyer & Sendlmeier, 1997). The measure *h1-h2* describes the amplitude difference between the first and second harmonics, which is the acoustic correlate for the relationship between the open and closed phases of the glottis, i.e., the space between vocal folds (Barsties V Latoszek, Maryn, et al., 2018).[4] If there are more open phases than closed phases, the value of h1-h2 will be higher, indicating a perceived breathiness. If more closed than open phases exist, the h1-h2 quotient will have negative values, indicating perceived creakiness (Barsties V Latoszek, Maryn, et al., 2018; Hillenbrand & Houde, 1996).

In addition to h1-h2, *spectral slope* and *tilt* are essential indicators of breathiness and creakiness in voice. Spectral slope is the rate of amplitude decrease between two increasing frequencies in a spectrum (McAleer et al., 2014).[5] It quantifies how signal amplitude diminishes with rising frequency. Spectral tilt describes the degree to which intensity drops off as frequency increases, i.e., the slope of the trend line through the spectrum (M. Gordon & Ladefoged, 2001). Both acoustical measures are typically negative values as they measure amplitude decreases. A breathier voice is normally associated with a relatively steep spectral slope and tilt, while a creakier voice is associated with a flat spectral slope and tilt (M. Gordon & Ladefoged, 2001; Kuang & Liberman, 2018). Additionally, speech is perceived as louder when spectral slope and tilt decrease, but this loudness is perceived as brightness and articulatory clarity, such as stressed vowels in speech, instead of an overall increase in loudness (Duvvuru & Erickson, 2013; Sluijter & van Heuven, 1996; Stévens & Hall, 1966).

Lastly, several studies have demonstrated that the *smoothed cepstral peak prominence (CPPS)* is another reliable indicator for dysphonic voices, particularly for breathy voices (Barsties V Latoszek et al., 2017; Hillenbrand & Houde, 1996; Pearsell & Pape, 2023). CPPS quantifies the prominence of the harmonic structure, i.e., periodic vocal fold vibrations, over noise in voice (Baker et al., 2022; Fraile & Godino-Llorente, 2014). Lower CPPS values are associated with higher perceived breathiness, whereas higher CPPS values indicate a “clear” voice.

[4] Harmonics are multiples of the fundamental frequency (f0) produced in the vocal folds. The 1st harmonic is the fundamental frequency (f0; Collins & Missing, 2003).
[5] A spectrum represents the distribution of intensity based on the frequency of an acoustic signal at a specific point in time.

Rough or hoarse voices are produced through irregular or desynchronized vibrations of the vocal folds accompanied by random fluctuations of the glottal pulse and a very high laryngeal tension (Barsties V Latoszek, Bodt de, et al., 2018; Dejonckere et al., 1993; Gobl & Chasaide, 2003). An example of a hoarse voice with desynchronized vocal fold vibrations is displayed in **Figure 3D**. The measure *harmonics-to-noise ratio (HNR)* represents the relationship between the periodic (harmonics) and aperiodic (noise) components of a speech. The lower the HNR values, the rougher and hoarse a voice is perceived (Anjos de Oliveira et al., 2020; McAleer et al., 2014). Jitter and shimmer measure acoustic irregularities in vocal folds, which might influence the perception of a rough or hoarse voice (Farrus et al., 2007; Frühholz & Belin, 2019; Wendahl, 1963). *Jitter* measures short-term perturbation of fundamental frequency (f0), and *shimmer* measures short-term amplitude perturbation (Clark et al., 2007; Gobl & Chasaide, 2003).

The definitions of the 15 acoustical measures, classified according to the four sound-wave dimensions, facilitate a more comprehensive understanding of the components that constitute a voice and the aspects considered in this study. Furthermore, the detailed explanations provided for these acoustical measures assist in interpreting the results with greater precision, particularly for those without a background in linguistics or phonetics.

3. Methodology and Data

3.1. Study Design and Sample

In pursuing the research objective to develop a BVP-Model to determine how brand personalities are perceived through voice, this study employs a correlational design and uses external judgments of personalities. Therefore, in this study, naïve listeners evaluated voices according to perceived brand personality traits via an online survey consisting of three parts.

The first part consisted of demographic and filter questions. Participants who were native German speakers with no hearing impairment or at least a compensated hearing impairment were accepted for the study. The second part comprised the sequential rating of two voices on a five-point scale (from 1 = ”does not apply at all” to 5 = “applies completely”) on the extent to which they associated the voice with the provided brand personality traits. Before this study, a preliminary test was conducted to determine the optimal number of voice ratings a participant could provide without losing focus, ensuring the highest quality of responses. The pre-test indicated that a maximum of two voices could be rated before participants demonstrated a decline in attention, concentration, and willingness for voice evaluations (for pre-test details, please refer to Appendix B). Finally, in the third part, the participants were asked to indicate the device they used to complete the questionnaire (speakers or headphones) and the background noise level they experienced while filling out the questionnaire. All participants gave informed consent before participating.

In total, 2,123 participants completed the questionnaire through a German online access panel provider. Before analyzing the data, participants who made insufficient effort to answer the questionnaire were removed (J. L. Huang et al., 2012; Leiner, 2019). Thus, 65 participants with missing answers, five participants who did not pass attention checks, eleven participants who had not played the voices, and six participants who took less than half the median response time to complete the questionnaire (median was 530s) were removed (Jandura, 2018; Leiner, 2019). After this initial data cleaning, the remaining data from 2,036 participants were split according to first and second voice ratings, resulting in 4,072 individual voice ratings. A closer look at the individual voice ratings still required removing 127 records, which showed low variance in the personality trait ratings based on their response pattern (variance ≤ .2 was used as the threshold). After the second data cleaning, 3,945 individual voice ratings from 2,000 participants remained (1,945 participants with two voice ratings and 55 participants with single voice ratings; $M_{age} = 51$; 53% female). For an overview of socio-demographic and survey-related sample characteristics, please see Appendix C.

3.2. Brand Personality Rating

For the brand personality perception rating, the German translations of 64 brand personality traits taken from the three BPS of Aaker (1997), Geuens et al. (2009), and Grohmann (2009) were used. The three BPS resulted in 66 traits in total, but since the translations of the two traits "down-to-earth" and "rugged" occurred twice, only 64 traits were used in the survey. See **Table 3** for an overview of the brand personality dimensions and traits of all three BPS in English and German.

Table 3. Overview of Brand Personality Traits Used in Survey

Authors	BPS Dimension	BPS Traits English	BPS Traits German	Source German Translation
Aaker (1997)	Competence	confident, corporate, hard-working, intelligent, leader, reliable, secure, successful, technical	zuversichtlich, integrative, hart, arbeitend, intelligent, führend, zuverlässig, sicher, erfolgreich, technisch	Aaker (2019)
	Excitement	contemporary, cool, daring, exciting, unique, imaginative, young, independent, trendy, spirited, unique, up-to-date	zeitgemäß, cool, gewagt, aufregend, einzigartig, fantasievoll, jung, unabhängig, modisch, temperamentvoll, modern	
	Ruggedness	masculine, outdoorsy, rugged, tough, western	männlich,naturverbunden, **robust**, zäh, westlich	
	Sincerity	cheerful, down-to-earth, family-oriented, friendly, honest, original, real, sentimental, sincere, small-town, wholesome	heiter, **bodenständig**, familienorientiert, freundlich, ehrlich, ursprünglich, echt, gefühlvoll, aufrichtig, kleinstädtisch, gesund	
	Sophistication	charming, feminine, glamorous, good-looking, smooth, upper class	charmant, weiblich, glamourös, gutaussehend, weich, vornehm	
Geuens et al. (2009)	Activity	active, dynamic, innovative	aktiv, dynamisch, innovativ	Geuens et al. (2009)*
	Aggressiveness	aggressive, bold	aggressiv, frech	
	Emotionality	romantic, sentimental	romantisch, sentimental	
	Responsibility	down-to-earth, responsible, stable	**bodenständig**, verantwortlich, stabil	
	Simplicity	ordinary, simple	gewöhnlich, einfach	
Grohmann (2009)	Feminity	expresses tender feelings, fragile, graceful, sensitive, sweet, tender	zeigt zärtliche Gefühle, empfindsam, anmutig, einfühlsam, herzig, zart	Lieven (2014)
	Masculinity	adventurous, aggressive, brave, daring, dominant, sturdy	abenteuerlustig, angriffslustig, tapfer, wagemutig, dominant, **robust**	

Notes: The four colored and bolded words are the two traits that occurred twice in the German translations. *Geuens et al. (2009) conducted a cross-cultural validation of their developed BPS in ten different countries, including Germany. The German translations were received from the authors upon request. BPS = brand personality scale.

3.3. Vocal Stimuli

Repositories of databases designed for managing language corpora were screened to identify voices that could be used as vocal stimuli. These databases included Clarin, the Database for Spoken German (DGD), the Institute for Natural Language Processing (IMS) of the University of Stuttgart, and the Hamburg Centre for Speech Corpora (HZSK) of the University of Hamburg.[6] Following a comprehensive analysis of potential corpora, voices from the Jena

[6] Clarin: clarin.eu/; DGD: dgd.ids-mannheim.de/; IMS: ims.uni-stuttgart.de/; HZSK: slm.uni-hamburg.de/hzsk/

Speaker Set (JESS) were selected as suitable vocal stimuli for this study (Zäske et al., 2020). The JESS is a corpus of voice recordings from 120 unfamiliar speakers, 61 male and 59 female, consisting of German vowels, syllables, read text, and semi-spontaneous speech.

In comparison to other existing German corpora, the JESS was chosen due to four reasons: 1) the corpus contained voice recordings with neutral content, without emotions or controversial topics that could influence personality perceptions; 2) all speakers were recorded under controlled conditions without background noises or other disruptions, making the extracted acoustic measures from voices comparable; 3) the corpus offered the highest number of female and male speakers with an appropriate age distribution; and 4) all speakers were native Germans, not foreigners speaking German, whose accent could influence personality judgments.

This study only used snippets of the semi-spontaneous speech recordings describing a farmyard scene. The research purpose required using semi-spontaneous recordings, as personality traits are best reflected when speakers speak spontaneously and have no behavioral constraints (Johnstone & Scherer, 2000). Moreover, out of the 120 voices, only 96 (47 females) were used, as 24 voices exhibited audible dialects, speech disorders, or articulation difficulties (e.g., lisp or stutter), which meant that no suitable coherent sentences could be found that could be used as a stimulus snippet. Dialects can be associated with specific social classes, affecting personality perception (Krauss et al., 2002). Four independent research assistants who are native German speakers conducted the dialect evaluation. Additionally, self-reported regional accent data from the JESS supplementary material (Table S3 in Zäske et al., 2019) were considered. The voice samples were cut into snippets that were, on average, 20.1 seconds long (range: 15.4s - 23.1s). Since it has been shown that personality judgments are made unconsciously after only a few seconds, an average sample length of 20.1 seconds seemed appropriate for making personality judgments without becoming too familiar with the voice (Aronovitch, 1976; McAleer & Belin, 2019; Ray, 1986). The participants rated each voice sample on average 41 times. For an overview of the selected and excluded voices from the JESS for this study, please refer to Appendix D. Please refer to Appendix A for access to the study's vocal stimulus set.

The JESS indicated a specific age distribution characteristic, as it distinguished "young" (18-25 years) and "old" speakers (60-81 years). Thus, voices from speakers between 26-59 years were missing. However, during the JESS development process, Zäske et al. (2020) requested that listeners evaluate the voices based on characteristics, including the perceived age of the speakers. They discovered that the age of young voices was overestimated by an average of seven years, while the age of older voices was underestimated by an average of nine years. As mentioned earlier, the paraverbal channel of human speech can provide listeners with a wealth of information about the speaker, including age estimations (J. K. Gordon et al., 2019). Voice parameters like speaking rate, pitch, or pausing behavior were found to influence the perceived age of a speaker (Harnsberger et al., 2008; Shipp et al., 1992; Skoog Waller & Eriksson, 2016). Since perceived age was more important than actual age in the present perceptual study, the specific age distribution within the JESS did not represent a disadvantage. The JESS included perceived middle-aged voices because of the described overestimation/ underestimation of the age of young/ old voices (see **Figure 4**).

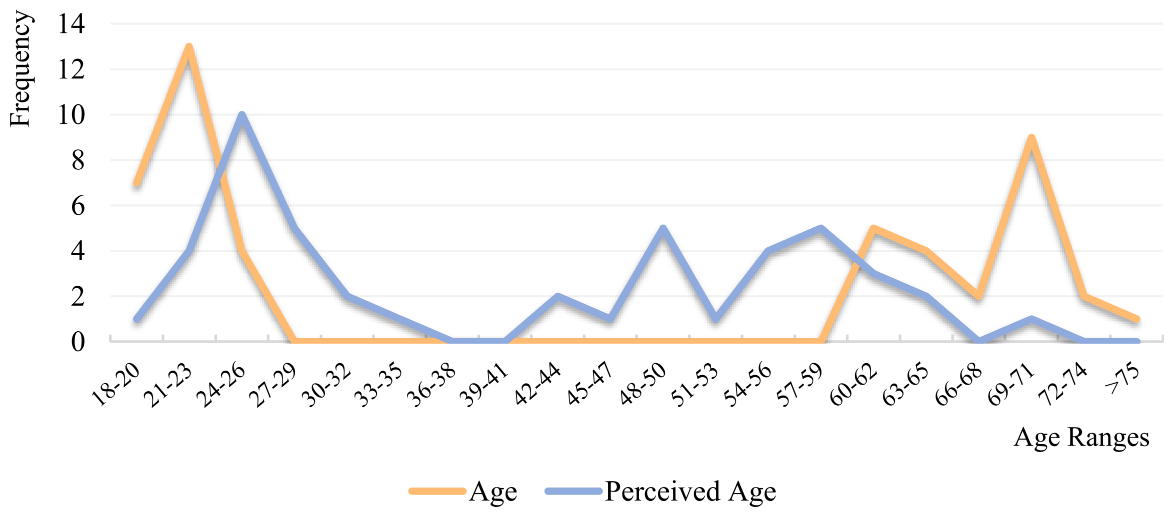


Figure 4. Age and Perceived Age of Speakers Within Vocal Stimuli
Notes: N = 96.

3.4. Acoustical Measures

The software PRAAT was used to extract the acoustical measures from the 96 voices (version 6.1.50; Boersma & Weenink, 1992-2022). In linguistics and communication sciences, this software tool for phonetic and speech analysis is widely used because of its user-friendly interface and publicly available extensions, plugins, and scripts. All JESS voice recordings were mono and had a sampling rate of 44.1kHz and a 16-bit resolution. They were Root Mean Square (RMS) normalized to 70dB (Zäske et al., 2020).

All acoustic-prosodic measures were calculated, and their numerical output was stored using a self-designed PRAAT script. Within the script, the Prosogram plugin (version 3.00f; Mertens, 2022) was used for extracting the data on articulation and speaking rate, silent pause duration, f0 mean, f0 SD, and f0 range; the Voiceprofile plugin (version 2.3.1; Mayer, 2019-2021) was used for extracting the data on all spectral features, i.e., h1-h2, spectral slope and tilt, HNR, CPPS, jitter, and shimmer. The intensity variability and silent pause frequency per minute were extracted through the analysis functions of PRAAT. A silent pause was defined as a pause that lasted at least .5 seconds without any audible sounds. The ranges and average values of the acoustic measures of the vocal stimuli set per gender are summarized in **Table 4**.

Table 4. Ranges and Means of Acoustic Measures of Vocal Stimuli Set "JESS"

	Female Voices		Male Voices	
Acoustical Measure (Metric)	**Range**	**Mean**	**Range**	**Mean**
Speaking Rate (syl/s)	1.58 - 4.79	3.08	2.07 - 4.64	3.31
Average Silent Pause Duration (s)	.52 - 2.66	1.40	.63 - 4.61	1.53
Silent Pause Frequency per Minute (n/m)	2 - 10	5.15	1 - 10	4.65
Articulation Rate (syl/s)	3.93 - 6.44	4.73	3.81 – 5.86	4.83
Intensity Variability (SD; dB)	9.66 - 15.88	12.68	9.91 - 17.48	12.49
Fundamental Frequency Mean (f0; Hz)	169.00 - 255.00	210.98	83.00 - 167.00	118.76
Fundamental Frequency Standard Deviation (SD f0; Hz)	24.91 - 84.89	50.09	11.75 - 40.89	24.89
Fundamental Frequency Range (f0-max – f0-min; st)	7.20 - 27.60	16.70	6.60 - 22.20	14.39
h1-h2 (dB)	(-1.81) - 7.21	2.31	(-3.59) – 5.57	.69
Spectral Slope	(-33.29) - (-18.77)	(-24.56)	(-32.01) - (-19.96)	(-25.05)
Spectral Tilt	(-12.16) - (-7.99)	(-10.89)	(-12.58) - (-9.80)	(-11.28)
Harmonics-to-Noise Ratio (HNR; dB)	13.52 - 24.17	19.51	10.34 - 19.04	14.18
Smoothed Cepstral Peak Prominence (CPPS; dB)	9.30 - 15.18	12.26	8.58 - 15.52	12.01
Jitter (%)	.96 - 3.33	1.91	1.58 - 3.98	2.89
Shimmer (%)	4.64 - 10.99	6.80	5.64 - 13.84	9.14

Notes: JESS = Jena Speaker Set; syl = syllable; s = second; m = minute; n = number; SD = standard deviation; Hz = Hertz; st = semitones; dB = Decibel.

4. Brand Voice Personalities

The first research question concerns identifying which brand personalities can be perceived through voice alone, i.e., brand voice personalities. As the items of three BPS are used to rate perceived brand personality through voice, it is first tested whether the scales can be confirmed within the data set. The confirmatory factor analyses (CFA) show that none of the three BPS can be reproduced. To achieve sufficient discriminant validity and an adequate model fit, excluding at least one dimension of each BPS being used would be necessary. This exclusion is required because the Fornell-Larcker criterion[7] is not met, and model fit indices such as the Tucker-Lewis Index (TLI) or Comparative Fit Index (CFI) do not meet the required threshold values (see Appendices E-G for the CFA results; Fornell & Larcker, 1981; Hu & Bentler, 1999). Thus, as the existing BPS cannot be replicated, certain predetermined brand personality traits could not be conveyed solely through the voice.

[7] Fornell-Larcker criterion: Average Variance Extracted (AVE) > squared correlation of the latent construct with the discriminant construct

As a result, an exploratory approach is chosen, and a scale is developed to indicate which brand personality traits and dimensions can be perceived through voice alone: the brand voice personality scale (BVP-Scale). For the scale development, the sample is split into two equivalent subsamples to perform an exploratory factor analysis (EFA) on the first subsample and a CFA on the second subsample. Statistical analyses are conducted using the software R (version 4.3.1) with the packages psych for EFA (version 2.4.3; Revelle, 2024), and lavaan for CFA (version 0.6-17; Rosseel, 2012). The four-step scale development and evaluation process is described in the following subsection and illustrated in **Figure 5**.

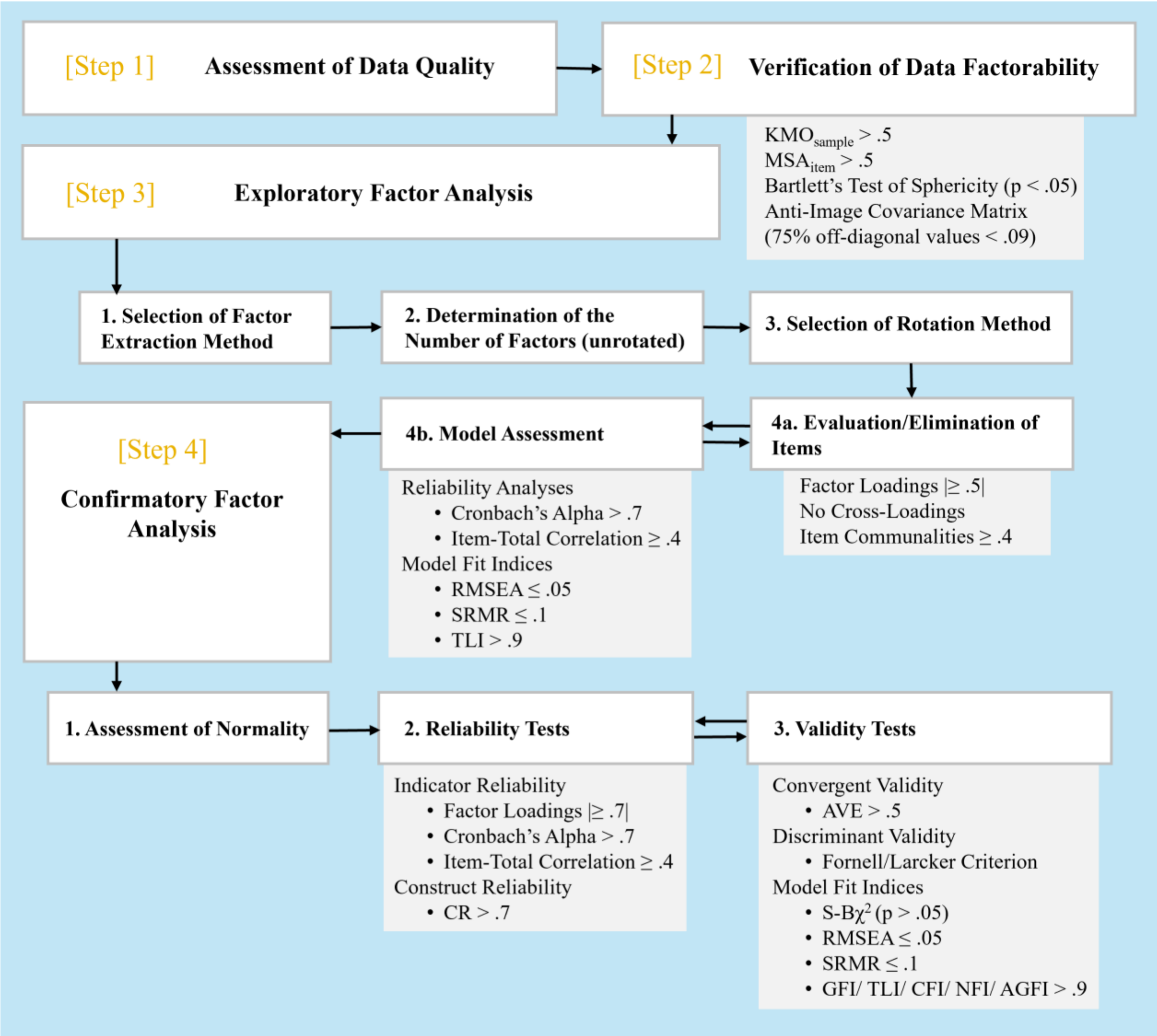


Figure 5. Four-Step Scale Development and Evaluation Process

Notes: Own illustration following the processes of S. Carpenter (2018) and Morgado et al. (2017).

Abbreviations: AGFI = Adjusted Goodness-of-Fit Index; AVE = Average Variance Extracted; CFI = Comparative Fit Index; CR = Composite Reliability; GFI = Goodness-of-Fit Index; KMO = Kaiser-Meyer-Olkin; MSA = Measure of Sampling Adequacy; NFI = Normed Fit Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; S-Bχ^2 = Satorra-Bentler Chi-squared; TLI = Tucker-Lewis Index.

4.1. Development of Brand Voice Personality Scale

4.1.1. Step 1: Assessment of Data Quality

First, the data quality is assessed by investigating whether there are any semantically illogical correlations between rated brand personality traits. Although no such correlations are found, it is observed that the ratings of the items "masculine" and "feminine" only referred to the gender of the voice. In other words, male voices are rated high on "masculine" (male voices: $M_{masculine} = 4.61$), and female voices are rated high on "feminine" (female voices: $M_{feminine} = 4.56$). Furthermore, there are no correlations between these and other items, except for a negative correlation of r = -.9 between them (see correlogram in Appendix H). This result suggests that male voices rated high on "masculine" are at the same time rated low on "feminine" (male voices: $M_{feminine} = 1.16$), female voices rated high on "feminine" are at the same time rated low on "masculine" (female voices: $M_{masculine} = 1.2$). These two gender-related personality traits are taken from Aaker's BPS (1997), which the researcher has been criticized for using, as they represent socio-demographic information instead of personality traits (Geuens et al., 2009). To focus on brand-related personality traits and since the results do not add value, the items "masculine" and "feminine" are excluded from further analyses, leaving 62 rated personality traits.[8]

4.1.2. Step 2: Verification of Data Factorability

The sample of 3,945 individual voice ratings is split into two equivalent subsamples ($n_{subsample1} = 1{,}970$; $n_{subsample2} = 1{,}975$) according to the Solomon method, a sample splitting technique developed explicitly for factor analyses. The split is successful as the similarity can be assessed through a communality ratio (S)[9] of 1, which indicates that both subsamples have

[8] An additional exploratory factor analysis of the personality traits assessed, including the two gender-related items "masculine" and "feminine", revealed that these two items form a separate dimension with opposite factor loadings, confirming that there are no correlations with other items and that they were used only to assess speaker's gender.

[9] $S = \frac{\min(KMO1, KMO2)}{\max(KMO1, KMO2)}$. KMO = Kaiser-Meyer-Olkin statistic

a similar amount of shared variance based on the Kaiser-Meyer-Olkin (KMO) statistic ($KMO_{subsample1}$ = .98; $KMO_{subsample2}$ = .98). (Lorenzo-Seva, 2021)

First, an EFA is conducted using subsample 1 (n = 1,970 individual voice ratings) to determine the dimensionality of the BVP-Scale. The suitability of the data for factor analysis is assessed using the KMO criterion together with the personality traits' Measure of Sampling Adequacy (MSA), Bartlett's Test of Sphericity, and the anti-image covariance matrix (Bartlett, 1950; Hair et al., 2013; Kaiser, 1974; Kaiser & Rice, 1974). The KMO measure of subsample 1, along with the MSA values of each personality trait, exceed the threshold of .5 (lowest $MSA_{s-imple}$ = .78). Additionally, Bartlett's Test of Sphericity yields a statistically significant result ($\chi^2(1891)$ = 71,367.46, p < .001), which tests the null hypothesis that the correlation matrix shows unrelated variables (Dziuban & Shirkey, 1974). Furthermore, the anti-image covariance matrix demonstrates that less than 25% of off-diagonal values exceed .09, indicating that the personality traits exhibit a high common variance and that correlations can be expected (Backhaus et al., 2021). Thus, all testing assumptions are met, indicating that the data is suitable for factor analysis.

4.1.3. Step 3: Exploratory Factor Analysis

Principal axis factoring (PAF) with oblique rotation (oblimin) with Kaiser-Normalization is chosen for the EFA. The PAF extraction method is employed as the purpose is to understand the latent constructs that account for the relationships among the measured variables as opposed to reducing the number of variables through principal component analysis (PCA; Backhaus et al., 2021). Additionally, it is assumed that the total amount of variance cannot be explained and that variables are linear combinations of latent constructs and residuals (Backhaus et al., 2021). This assumption is based on the fact that the original goal of this study was not to develop a scale for measuring brand voice personalities. Instead, it arose due to the inability to confirm any existing BPS. As a result, the generated item pool contains only some of the brand personalities relevant to voice perception. Therefore, PCA is not an appropriate extraction method because it assumes that variables are linear combinations of latent constructs without residuals, attempting to explain the total variance. An oblique rotation is chosen because correlations between constructs in personality research are assumed (in contrast to an orthogonal rotation), and the oblimin rotation is selected due to the best model fit (DeVellis, 2017).

The scree plot test without factor rotation suggests a 5-factor model based on the Kaiser criterion (eigenvalues greater than 1.0; see Appendix I; Cattell, 1966; Zwick & Velicer, 1986). Analysis of the oblique rotated factors suggests a 4-factor model with items that had factor loadings greater than or equal to |.5| (Hair et al., 2020). At this stage, the model is purified, resulting in the elimination of 30 personality traits with factor loadings lower than |.5| and communalities lower than .4. This purification leaves a 4-factor model with 32 items for further analyses (see Appendix J).

The model assessment shows a good model fit with a Root Mean Square Error of Approximation (RMSEA) of .04 (≤ .05 for good model fit), a Standardized Root Mean Square Residual (SRMR) of .02 (≤ .1 for good model fit), and a Tucker-Lewis Index (TLI) of .95 (>.9 for good model fit; Backhaus et al., 2015; Hu & Bentler, 1999). Nonetheless, the fourth construct's low Cronbach's alpha = .65 with the two items "young" and "up-to-date" indicate low internal consistency (Cronbach's alpha should exceed .7), which is why it is decided to eliminate this factor with these two items, leaving a 3-factor model with 30 items for further analysis.

Analyses of the new rotated 3-factor model with 30 items suggest further eliminating the items "rugged" due to a factor loading lower than |.5| and "down-to-earth" due to a communality lower than .4. The final EFA result, a 3-factor model with 28 items shows factor loadings greater than |.5|, no cross-loadings, and communalities greater than .4 for all constructs. Further, the model shows high reliability (Cronbach's alpha > .7; item-total correlations ≥ .4) and good model fit (SRMR = .02; RMSEA = .04; TLI = .95). See Appendix K for the final EFA result.

4.1.4. Step 4: Confirmatory Factor Analysis

To validate the BVP-Scale, a series of CFAs are conducted with the second subsample (n = 1,975 individual voice ratings). The univariate and multivariate normality assessment indicates that the data does not follow a normal distribution (Weiber & Mühlhaus, 2014; West et al., 1995; see Appendix L). Thus, in the CFAs, the maximum likelihood estimation with robust standard errors and the Satorra-Bentler (S-B) correction is used to consider the non-normality of the data (Hu & Bentler, 1999; Tian et al., 2001). In the iterative process of six CFAs, 17 items are eliminated based on the application of criteria such as low factor loadings or item communalities. This process was continued until the resulting model exhibited sufficient validity and reliability and demonstrated an acceptable fit. This iterative process, in which validity and reliability tests were carried out at the indicator, construct, and model levels, is described in the following.

Indicator reliability is given when factor loadings are greater than or equal to |.7| and are significant (p-value < .5/ Critical Ratio (C.R.) > 1.96; Backhaus et al., 2015). In addition, item-total correlations greater than or equal to .4 show a strong relation to the construct or scale, and Cronbach's alpha greater than .7 indicates high **internal consistency** of the item set within the scale (Cronbach, 1951; Hair et al., 2013). In the first and second CFA, 13 items[10] showed low indicator reliability due to factor loadings lower than |.7| and were eliminated.

Construct reliability and **internal consistency** are achieved when composite reliability[11] (CR) exceeds .7 (Hair et al., 2013). While both Cronbach's alpha and CR measure

---

[10] aggressive (from Grohman, 2009), daring (from Aaker, 1997), dominant, glamorous, graceful, imaginative, leader, real, responsible, secure, sentimental (from Geuens et al., 2009), stable, wholesome

[11] $Composite\ Reliability\ (CR) = \frac{(\sum \text{Standardized Factor Loadings})^2}{(\sum \text{Standardized Factor Loadings})^2 + \sum Measurement\ Error\ Variances}$

reliability and internal consistency, it is crucial to examine both measures. Cronbach's alpha, a traditional measure of internal consistency, may underestimate reliability due to the assumption of equal factor loadings and equal error variances. In contrast, CR considers the actual factor loadings of each item, potentially providing a more accurate estimate of reliability (Hair et al., 2013). In the final CFA, both measures exhibited identical values, indicating that the measured items representing a construct are highly reliable and consistent.

Convergent validity describes the extent to which items of a construct share a high proportion of variance in common and can be measured through the Average Variance Extracted[12] (AVE; Hair et al., 2013). For adequate convergent validity, the AVE should be greater than .5, which all constructs achieved (Backhaus et al., 2015; Hair et al., 2013). **Discriminant validity** refers to the degree to which a construct is distinct from other constructs and can be measured using the Fornell-Larcker criterion. According to the criterion, the construct's AVE should exceed the squared correlation between that construct and other constructs in the model (Fornell & Larcker, 1981). The final model met this requirement (see **Table 5**). **Overall construct validity** is achieved when the model reliably measures what it is intended to measure, with high internal consistency on both indicator and construct levels and proven convergent and discriminant validity (Hair et al., 2013). Based on the described thresholds, the final model demonstrated high construct validity.

Table 5. Average Variance Extracted and Squared Correlations

	Sincerity	**Sensitivity**	**Excitement**
Sincerity	(.60)		
Sensitivity	.48	(.66)	
Excitement	.18	.23	(.56)

Notes: Average Variance Extracted (AVE) in brackets.

At the model level, the observed covariance matrix and the estimated covariance matrix are compared to assess the **model fit** (Backhaus et al., 2015). Several fit statistics were analyzed

[12] $Average\ Variance\ Extracted\ (AVE) = \frac{\sum(\text{Standardized Factor Loadings})^2}{\sum(\text{Standardized Factor Loadings})^2 + \sum Measurement\ Error\ Variances}$

to evaluate the goodness-of-fit described in the following (Backhaus et al., 2015; Hair et al., 2013).

The Chi-square test (S-Bχ^2) was 146.971 with 41 degrees of freedom (df). The associated p-value of .001 was significant, indicating that the observed covariance matrix did not match the estimated covariance matrix within the sampling variance. The measure of Chi-square is directly related to the sample size. Due to the large data set of 1,975 individual voice ratings in subsample 2, this measure could not be meaningfully used to evaluate the model's fit to the data (Hair et al., 2013; Homburg et al., 2015). Thus, further model fit statistics were closely examined.

The *absolute fit measures*, which assess how well the theoretical model fits the sample data, Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Residual (SRMR), Goodness-of-Fit (GFI), and the Adjusted Goodness-of-Fit Index (AGFI) provided support for a good model fit with RMSEA = .04 (≤ .05 for good model fit), SRMR = .03 (≤ .1 for good model fit), GFI = .98, and AGFI = .97 (both > .9 for good model fit; Hair et al., 2013; Hu & Bentler, 1999). Further, Normed Fit Index (NFI), Tucker-Lewis Index (TLI), and Comparative Fit Index (CFI) were all higher than the proposed threshold of .9 for a good model fit (NFI = .98, TLI = .98, CFI = .99; Hair et al., 2013; Hu & Bentler, 1999). These *incremental/comparative fit measures* assess the proportion improvement in fit by comparing the model of interest with a baseline model (null model; Hu & Bentler, 1999).

While 13 items were eliminated in the first two CFAs due to low factor loadings, the following CFA revealed high reliability and validity at the indicator and construct level but not at the model level. Four additional CFAs were conducted, with one item iteratively removed[13] until the model fit was reasonable based on the described fit indices. These purifications led to the final CFA - a reflective first-order model with three brand voice personality dimensions

[13] tender, expresses tender feelings, sweet, romantic

(constructs) and eleven brand voice personality traits (items; see **Table 6**). The dimensions are labeled and defined as follows:

1) **Sincerity**: This dimension describes a brand voice personality perceived as honest, reliable, and sincere.
2) **Sensitivity**: This dimension describes a brand voice personality perceived as smooth, fragile, sentimental, and sensitive.
3) **Excitement**: This dimension describes a brand voice personality that is perceived as spirited, adventurous, daring, and exciting.

Table 6. Final Confirmatory Factor Analysis

Construct/ Item	Standardized Factor Loading (≥ |.7|)	Item-Total Correlation (≥ .4)	CR (> .7)	AVE (> .5)	Fornell-Lacker Criterion
Sincerity (α = .82)					
sincere	.82***	.77	.82	.60	Yes
honest	.75***	.73			
reliable	.75***	.72			
Sensitivity (α = .88)					
sensitive	.87***	.83	.88	.66	Yes
sentimental (Aaker, 1997)	.86***	.84			
fragile	.79***	.74			
smooth	.72***	.72			
Excitement (α = .83)					
adventurous	.76***	.74	.83	.56	Yes
exciting	.75***	.73			
spirited	.75***	.73			
daring (Grohman, 2009)	.72***	.71			

Notes: Method: Maximum likelihood estimation with robust standard errors and Satorra-Bentler (S-B) correction. n = 1,975 individual voice ratings; ***p < .001. Fornell-Lacker criterion: AVE > squared correlations between the constructs.
Global Model Fit Indices: S-Bχ^2 (41) = 146.971 (p < .001); χ^2/df = 4.3; Standardized Root Mean Square Residual (SRMR) = .03; Root Mean Square Error of Approximation (RMSEA) = .04; Comparative Fit Index (CFI) = .99; Tucker-Lewis Index (TLI) = .98; Goodness-of-Fit Index (GFI) = .98; Adjusted Goodness of Fit Index (AGFI) = .97; Normed Fit Index (NFI) = .98.
Abbreviations: CR = Composite Reliability; AVE = Average Variance Extracted.

Finally, it is investigated whether an alternative model exists that achieves a better BVP-Scale model due to a different dimensionality (Brakus et al., 2009; Tian et al., 2001). Therefore, the following models are compared: a *null model*, assuming no correlations between items; a *1-factor model*, considering all items loaded on a single construct; a *2-factor model*, considering the dimensions sincerity and sensitivity are one construct as these two constructs showed the highest correlation of .69; and a *3-factor model*, assuming the constructs the final CFA proposed. The results support the 3-factor model with the brand voice personality dimensions of sincerity, sensitivity, and excitement (see **Table 7**)

Table 7. Model Fit Indices for Competing Measurement Models

Competing Models	S-Bχ^2	df	CFI	TLI	RMSEA	SRMR	AIC
Null Model	9,091.337	55	NA	NA	NA	NA	NA
One-factor Model	2,488.607	44	.73	.66	.17	.13	58,760.107
Two-factor Model	895.390	43	.91	.88	.1	.06	56,750.818
Three-factor Model	**146.971**	**41**	**.99**	**.98**	**.04**	**.03**	**55,852.157**

Notes: The results highlighted in bold are considered the best.
Abbreviations: AIC = Akaike Information Criterion; CFI = Comparative Fit Index; df = degrees of freedom; NA = not applicable; RMSEA = Root Mean Square Error of Approximation; S-Bχ^2 = Satorra-Bentler Chi-squared; SRMR = Standardized Root Mean Square Residual; TLI = Tucker-Lewis Index.

4.1. Interpretation of Brand Voice Personality Scale

To achieve the study's objective of developing a model that determines how brand personalities are perceived through voice, this study first identified which brand personalities can be perceived through voice alone. The survey's data analyses, in which naïve listeners were asked to rate voices according to brand personalities, revealed three brand personality dimensions that can be conveyed through voice alone: sincerity, sensitivity, and excitement. These personality dimensions can be measured with the developed brand voice personality scale (BVP-Scale) consisting of eleven brand voice personality traits (see **Figure 6**). The brand voice personality dimension sincerity describes an honest, reliable, and sincere brand. The sensitivity personality dimension describes a brand voice perceived as smooth, fragile, sentimental, and sensitive. Finally, a spirited, adventurous, daring, and exciting brand is expressed through the brand voice personality dimension of excitement.

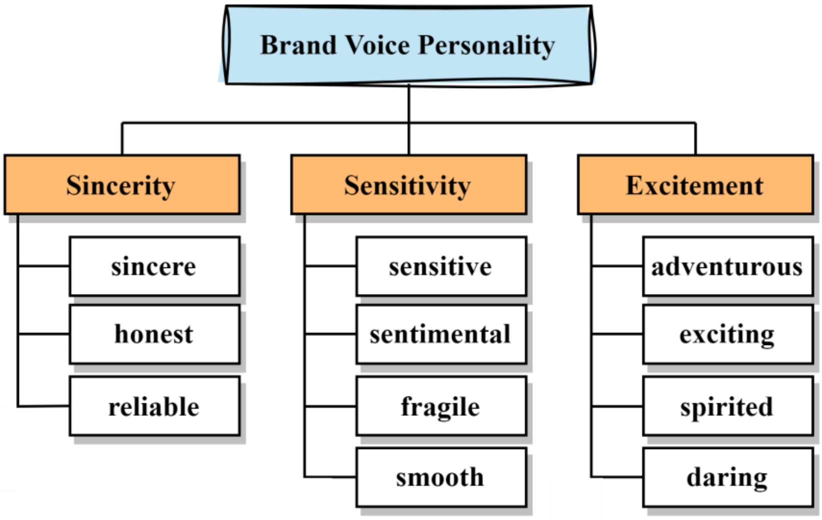


Figure 6. Brand Voice Personality Scale (BVP-Scale)

The findings follow prior research on the identification of (human) personalities through voice by Scherer (1972), Zuckerman and Driver (1988), and McAleer et al. (2014). These studies concluded that only two to three key traits typically derive from the voice, and all other personality traits are associated in "halo clusters" (McAleer & Belin, 2019, p. 589). For instance, listeners may initially perceive a speaker as dominant due to a low-pitched voice. In addition, the same speaker can be associated as competent or aggressive, depending on which personality traits the listener associates with dominance. Consequently, listeners are more likely to perceive a limited number of key traits combined in halo clusters, which form a personality dimension, as opposed to perceiving many distinct personality dimensions. In this study, three halo clusters are identified, which are represented by the derived three brand voice personality dimensions.

This study is the first to investigate the vocal perception of brand personalities. Its results show that commonly known dimensions describing brand personalities, such as those identified by Aaker (1997), Geuens et al. (2009), and Grohman (2009), cannot all be perceived through voice. For the development of BPS, typically well-known brands (e.g., Coca-Cola, Mercedes, Apple) were presented to participants, who thought of the brand as if it were a person and assigned personality traits to it (Aaker, 1997). The participants evaluated the brand personalities on the general elements of brand communication, including visual elements like a brand logo, shapes, and colors, and their experiences with the brand in mind. Consequently, the derived brand personalities of BPS were based on multimodal perceptions. However, if any information is missing, the perception of a personality is formed based on the remaining cues. This is the case in the auditory communication between the brand and the consumer using voice AI, where visual information is typically absent (Packard & Berger, 2024). This study has shown that the exclusive presence of acoustical cues leads to the perception of the brand voice personality dimensions of sincerity, sensitivity, and excitement, which can be measured through the presented BVP-Scale.

5. Vocal Perception of Brand Voice Personalities

The second research question concerns determining which combinations of voice parameters in female and male voices induce perceptions of brand voice personalities, forming the BVP-Model. Thus, it is investigated how the 15 acoustical measures (i.e., exogenous latent variables; described in section 2.2.1) relate to the three brand voice personality dimensions sincerity, sensitivity, and excitement (i.e., endogenous latent variables) of the developed BVP-Scale. Therefore, structural equation modeling (SEM) is applied. Participants' socio-demographic data (gender, age, profession, education, income) and study-related data (background noise, device used) are included in the model as covariates.

5.1. Development of Brand Voice Personality Model

5.1.1. Multicollinearity Assessment

A requirement for SEM is the reduction of multicollinearity between exogenous latent variables (Westlund et al., 2008). Correlations are expected between timing measures such as speaking rate, average silent pause duration, and silent pause frequency per minute, as these acoustical measures describe pausing behavior in speech. Also, correlations are expected between several voice quality measures like HNR, jitter, and shimmer, as these parameters measure perturbations in voice. Therefore, the first step is to examine the pairwise correlation coefficients between the 15 acoustical measures (see Appendix M; Hair et al., 2020; Sarstedt & Mooi, 2019). Five pairs of correlation coefficients show values exceeding |.8|, indicating a strong correlation, thereby suggesting a multicollinearity problem.[14] As HNR shows strong correlations with three other acoustical measures (f0 mean, jitter, and shimmer), it is decided to eliminate this acoustical measure from further analysis.

The strong correlations between f0 SD and f0 mean, as well as jitter and shimmer, are further examined by computing each item's variance inflation factor (VIF). In doing so, a multiple regression is run for each acoustic measure on all other acoustic measures. Higher VIF values imply that the variance of an acoustical measure can be explained by the other acoustical measures in the model, which indicates (multi)collinearity (Sarstedt et al., 2020). Critical collinearity issues likely occur if VIF values are greater than 5, which is used as a threshold (Backhaus et al., 2021; Sarstedt & Mooi, 2019). Seven acoustical measures show VIF values above 5, with silent pause frequency per minute (VIF = 11.71), f0 SD (VIF = 10.93), and f0 mean (VIF = 10.91) showing the highest values (see Appendix N). Acoustic measures with the highest VIFs are gradually eliminated until multicollinearity is no longer critical. The three acoustical measures silent pause frequency per minute, f0 SD, and jitter are eliminated from the model in this process. Finally, eleven acoustical measures show acceptable VIF values, which can be used for further SEM computation (see **Table 8**).

[14] Pairwise correlation coefficients greater than |.8|: f0 SD and f0 mean (.82), HNR and f0 mean (.85), jitter and shimmer (.87), HNR and shimmer (-.88), and HNR and jitter (-.90).

Table 8. Conclusive Multicollinearity Assessment

Acoustical Measures	R^2	Collinearity Statistics	
		Tolerance > .2	VIF < 5
Speaking Rate	.45	.55	1.80
Average Silent Pause Duration	.36	.64	1.57
Articulation Rate	.40	.60	1.66
Intensity Variability	.53	.47	2.11
f0 Mean	.70	.30	3.33
f0 Range	.33	.67	1.49
h1-h2	.31	.69	1.44
Spectral Slope	.50	.50	1.98
Spectral Tilt	.32	.68	1.46
CPPS	.72	.28	3.56
Shimmer	.77	.23	4.27

Notes: Tolerance = 1 – R^2; VIF = 1/(1- R^2). CPPS = smoothed cepstral peak prominence; f0 = fundamental frequency; SD = standard deviation; VIF = variance inflation factor.

5.1.2. Measurement Invariance Testing

Since vocal personality perception depends on the speaker's gender, the data from male and female voices are examined separately in this study. Thus, a multigroup SEM with a group code approach is conducted to estimate the speaker's gender-specific effects on vocal brand personality perception. To compare mean differences between groups, it is required to test for configural (number of factors and the pattern of factor loadings are equivalent across groups), metric (magnitude of factor loadings is equivalent across groups), and scalar (scale intercepts are equal across groups) invariance (Steenkamp & Baumgartner, 1998). Measurement invariance tests show significant differences between the metric and scalar models, indicating that scalar invariance is not supported (see **Table 9**). Therefore, the anticipated difference in the brand personality perception between female and male voices is confirmed, as the values of the latent means across the two genders cannot be compared and must be observed separately.

Table 9. Measurement Invariance Testing

Invariance Test	S-Bχ^2	df	RMSEA	CFI	ΔS-Bχ^2	Δdf	ΔRMSEA	ΔCFI	p
Configural Invariance	908.567	368	.029	.971					
Metric Invariance	916.654	376	.029	.971	8.087	8	.000	.000	.41
Scalar Invariance	1,002.049	384	.030	.967	85.395	8	.001	.004	<.001

Notes: Δ = difference in respective estimates across two nested models. CFI = Comparative Fit Index; df = degrees of freedom; RMSEA = Root Mean Square Error of Approximation; S-Bχ^2 = Satorra-Bentler chi-squared.

5.1.3. Structural Equation Model

The structural model, according to the conceptual framework with group-specific effects, shows a good model fit (S-Bχ^2(368) = 908.567, $p < .001$; RMSEA = .029; SRMR = .017; CFI = .971, TLI = .961; Hu & Bentler, 1999). Here, the covariance between the traits "honest" and "sincere" is allowed within the brand voice personality dimension sincerity. This adjustment is indicated by the corresponding modification indices, which is plausible due to their similar meanings in German. **Tables 10-12** show the estimation results of the SEM for female and male voices per brand voice personality dimension sincerity, sensitivity, and excitement. The SEM computation is conducted using the software R (version 4.3.1) with the package lavaan (version 0.6-17; Rosseel, 2012).

For the perception of the **brand voice personality dimension sincerity** the structural model shows significant correlations for *female voices* with speaking rate ($\beta = .12$, $z = 2.67$, $p < .05$), h1-h2 ($\beta = .12$, $z = 3.56$, $p < .001$), spectral slope ($\beta = .09$, $z = 2.12$, $p < .05$), and CPPS ($\beta = -.12$, $z = -2.31$, $p < .05$; see **Table 10A**). The positive correlation with speaking rate suggests that a fluent speech pattern relates to the perception of a sincere brand. Furthermore, the positive correlation with h1-h2 and the negative correlation with CPPS indicates that a breathy voice induces perceptions of sincerity. The positive correlation with spectral slope suggests that creakiness and brightness relate to sincere brands.

For the perception of the **brand voice personality dimension sincerity**, the structural model shows significant correlations for *male voices* with speaking rate ($\beta = .12$, $z = 2.22$, $p < .05$), spectral slope ($\beta = .07$, $z = 1.84$, $p < .1$), and spectral tilt ($\beta = .07$, $z = 2.18$, $p < .05$; see **Table 10B**). Furthermore, the listener's gender negatively correlates with sincerity perceptions in male voices ($\beta = -.08$, $z = -3.02$, $p < .05$), which means that sincerity in male voices is perceived when the listener is female. Apart from the listener's gender influence on the perception of a sincere brand, the positive correlation with speaking rate suggests that a fluent speech pattern relates to sincere brand representation when using male voices. Furthermore, the positive correlations with spectral slope and tilt indicate that creakiness and brightness in voice induce sincerity perceptions.

Table 10. Estimated Regression Coefficients of the Structural Equation Models for Brand Voice Personality Sincerity

		A. Female Voices $R^2 = .036$			B. Male Voices $R^2 = .025$		
Relationship	**Path**	**β**	**S.E.**	**z-value**	**β**	**S.E.**	**z-value**
Acoustical Measures → Sincerity	Speaking Rate → Sincerity	.12	.06	2.67**	.12	.09	2.22**
	Av. Pause Duration → Sincerity	.03	.08	.56	-.03	.04	-1.08
	Articulation Rate → Sincerity	-.05	.05	-1.37	-.06	.10	-1.22
	Intensity Variability → Sincerity	.00	.03	.04	.03	.03	.69
	f0 Mean → Sincerity	-.02	.01	-.71	-.01	.02	-.33
	f0 Range → Sincerity	.01	.01	.24	.02	.01	.76
	h1-h2 → Sincerity	.12	.01	3.56***	.01	.01	.19
	Spectral Slope → Sincerity	.09	.01	2.12**	.07	.01	1.84*
	Spectral Tilt → Sincerity	-.05	.04	-1.40	.07	.04	2.18**
	CPPS → Sincerity	-.12	.03	-2.31**	-.09	.03	-1.46
	Shimmer → Sincerity	-.03	.03	-.53	-.04	.03	-.84
Covariates → Sincerity	Listener's Gender[a] → Sincerity	.01	.04	.37	-.08	.05	-3.02**
	Listener's Age → Sincerity	-.02	.01	-.55	-.02	.02	-.81
	Listener's Education → Sincerity	.02	.01	.71	.00	.01	.00
	Listener's Profession → Sincerity	-.01	.01	-.38	.03	.01	.99
	Listener's Income → Sincerity	-.04	.03	-1.35	-.02	.03	-.84
	Background Noise → Sincerity	.02	.06	.58	.02	.06	.84
	Device Used[b] → Sincerity	-.02	.05	-.78	-.01	.05	-.53

Notes: Estimates represent standardized path coefficients. Significant paths have been highlighted in orange to improve orientation. The darker the color, the higher the significance level. [a] dummy coded: "0" = female, "1" = male; [b] dummy coded: "0" = speaker, "1" = headphones; *p < .1; ** p < .05; ***p < .001; av. = average; S.E. = standard errors.

For the perception of the **brand voice personality dimension sensitivity** the structural model shows significant correlations for *female voices* with average pause duration (β = -.03, z = -1.12, p < .1), articulation rate (β = -.04, z = -.94, p < .1), intensity variability (β = -.07, z = -1.61, p < .05), f0 mean (β = -.01, z = -.29, p < .1), h1-h2 (β = .06, z = 1.95, p < .001), CPPS (β = -.12, z = -2.25, p < .001), and shimmer (β = -.09, z = -2.03 p < .05; see **Table 11A**). The negative correlations with average pause duration, articulation rate, and intensity variability suggest that a slower, less dynamic, but more fluent speech pattern with short silent pauses relates to the perception of a sensitive brand. Furthermore, female voices representing a brand with a lowered pitch are likelier to be associated with sensitivity than those with a higher pitch.

In addition, the positive correlation with h1-h2 and the negative correlation with CPPS indicate that a breathy voice induces sensitivity assessments. Roughness in the voice leads to decreased sensitivity perceptions, as indicated by the negative correlation with shimmer. Several significant correlations between the listener's gender ($\beta = .10$, $z = 4.17$, $p < .001$), age ($\beta = .06$, $z = 2.36$, $p < .05$), and income ($\beta = -.06$, $z = -2.18$, $p < .05$) and sensitivity perceptions exist. Accordingly, sensitivity is perceived in female voices when the listener is male, older, or has less income.

For the perception of the **brand voice personality dimension sensitivity**, the structural model shows significant correlations for *male voices* with h1-h2 ($\beta = .06$, $z = 1.95$, $p < .05$), spectral tilt ($\beta = .06$, $z = 2.09$, $p < .05$), CPPS ($\beta = -.12$, $z = -2.25$, $p < .05$), and shimmer ($\beta = -.09$, $z = -2.03$, $p < .05$; see **Table 11B**). The positive correlation with h1-h2 and the negative correlation with CPPS suggest that a breathy voice relates to sensitivity perceptions. The positive correlation with spectral tilt indicates that the voice should be creaky and bright, and a negative correlation with shimmer indicates that roughness in the voice leads to decreased sensitivity perceptions. Furthermore, the gender of the listener is found to positively correlate with sensitivity perceptions in male voices ($\beta = .05$, $z = 1.82$, $p < .1$). In contrast, the income of the listener is found to negatively correlate ($\beta = -.05$, $z = -1.96$, $p < .1$). These findings suggest that specifically when the listener is male or has less income, sensitivity perceptions are induced in male voices.

Table 11. Estimated Regression Coefficients of the Structural Equation Models for Brand Voice Personality Sensitivity

		A. Female Voices $R^2 = .080$			B. Male Voices $R^2 = .044$		
Relationship	**Path**	**β**	**S.E.**	**z-value**	**β**	**S.E.**	**z-value**
Acoustical Measures → Sensitivity	Speaking Rate → Sensitivity	.07	.10	1.39	.07	.10	1.39
	Av. Pause Duration → Sensitivity	-.03	.04	-1.12*	-.03	.04	-1.12
	Articulation Rate → Sensitivity	-.04	.11	-.94*	-.04	.11	-.94
	Intensity Variability → Sensitivity	-.07	.03	-1.61**	-.07	.03	-1.61
	f0 Mean → Sensitivity	-.01	.02	-.29*	-.01	.02	-.29
	f0 Range → Sensitivity	.03	.01	.80	.03	.01	.80
	h1-h2 → Sensitivity	.06	.02	1.95***	.06	.02	1.95**
	Spectral Slope → Sensitivity	.05	.01	1.46	.05	.01	1.46
	Spectral Tilt → Sensitivity	.06	.04	2.09	.06	.04	2.09**
	CPPS → Sensitivity	-.12	.04	-2.25***	-.12	.04	-2.25**
	Shimmer → Sensitivity	-.09	.03	-2.03**	-.09	.03	-2.03**
Covariates → Sensitivity	Listener's Gender[a] → Sensitivity	.10	.05	4.17***	.05	.05	1.82*
	Listener's Age → Sensitivity	.06	.02	2.36**	.00	.02	-.11
	Listener's Education → Sensitivity	-.04	.01	-1.48	-.01	.01	-.34
	Listener's Profession → Sensitivity	-.02	.01	-.57	-.01	.01	-.34
	Listener's Income → Sensitivity	-.06	.03	-2.18**	-.05	.03	-1.96*
	Background Noise → Sensitivity	.01	.07	.33	.03	.07	1.09
	Device Used[b] → Sensitivity	.01	.06	.50	.01	.06	.49

Notes: Estimates represent standardized path coefficients. Significant paths have been highlighted in orange to improve orientation. The darker the color, the higher the significance level. [a] dummy coded: "0" = female, "1" = male; [b] dummy coded: "0" = speaker, "1" = headphones; *p < .1; ** p < .05; ***p < .001; av. = average; S.E. = standard errors.

For the perception of the **brand voice personality dimension excitement** the structural model shows significant correlations for *female voices* with speaking rate (β = .20, z = 4.46, p < .001), average pause duration (β = .09, z = 1.98, p < .05), h1-h2 (β = .07, z = 2.29, p < .05), spectral slope (β = .14, z = 3.30, p < .001), and shimmer (β = -.09, z = -2.01, p < .05; see **Table 12A**). The positive correlations with speaking rate and average pause duration suggest that a fluent speech pattern with longer pauses leads to the perception of an exciting brand. These findings are contradictory since the length of pauses typically disrupts speech fluency. Moreover, the positive correlations with h1-h2 and spectral slope indicate that a breathy, creaky, and bright voice induces excitement perceptions. Moreover, the negative correlation with shimmer suggests that the female voice representing an exciting brand should not be rough.

For the perception of the **brand voice personality dimension excitement**, the structural model shows significant correlations for *male voices* with average pause duration ($\beta = .05$, $z = 1.66$, $p < .1$), spectral slope ($\beta = .14$, $z = 3.56$, $p < .001$), CPPS ($\beta = -.15$, $z = -2.63$, $p < .05$), and shimmer ($\beta = -.07$, $z = -1.65$, $p < .1$; see **Table 12B**). The positive correlation with average pause duration suggests that a speech pattern with longer pauses relates to perceiving a speaker as exciting. Furthermore, the positive correlation with spectral slope and the negative correlation with CPPS indicate that a breathy, creaky, and bright voice induces excitement assessments. Moreover, the negative correlation with shimmer suggests that the male voice representing an exciting brand should not be rough.

For both genders, there are several significant correlations between the listener's gender (female: $\beta = .11$, $z = 4.26$, $p < .001$; male: $\beta = .08$, $z = 3.02$, $p < .05$), age (female: $\beta = -.23$, $z = -8.38$, $p < .001$; male: $\beta = -.22$, $z = -8.57$, $p < .001$), and education (female: $\beta = -.08$, $z = -2.85$, $p < .05$; male: $\beta = -.07$, $z = -2.63$, $p < .05$) with excitement perceptions (see **Table 12**). Accordingly, female and male brand excitement perceptions increase when listeners are male, younger, or less educated.

Table 12. Estimated Regression Coefficients of the Structural Equation Models for Brand Voice Personality Excitement

Relationship	Path	A. Female Voices $R^2 = .101$			B. Male Voices $R^2 = .099$		
		β	S.E.	z-value	β	S.E.	z-value
Acoustical Measures → Excitement	Speaking Rate → Excitement	.20	.05	4.46***	.07	.08	1.31
	Av. Pause Duration → Excitement	.09	.07	1.98**	.05	.04	1.66*
	Articulation Rate → Excitement	.01	.05	.34	.06	.09	1.16
	Intensity Variability → Excitement	.04	.02	.89	-.03	.02	-.72
	f0 Mean → Excitement	-.02	.01	-.50	.01	.01	.28
	f0 Range → Excitement	.01	.01	.23	-.03	.01	-1.10
	h1-h2 → Excitement	.07	.01	2.29**	-.02	.01	-.71
	Spectral Slope → Excitement	.14	.01	3.30***	.14	.01	3.56***
	Spectral Tilt → Excitement	-.01	.03	-.36	.00	.04	.02
	CPPS → Excitement	-.08	.03	-1.62	-.15	.03	-2.63**
	Shimmer → Excitement	-.09	.02	-2.01**	-.07	.02	-1.65*
Covariates → Excitement	Listener's Gender[a] → Excitement	.11	.04	4.26***	.08	.04	3.02**
	Listener's Age → Excitement	-.23	.01	-8.38***	-.22	.01	-8.57***
	Listener's Education → Excitement	-.08	.01	-2.85**	-.07	.01	-2.63**
	Listener's Profession → Excitement	-.03	.01	-.93	.03	.01	1.18
	Listener's Income → Excitement	.00	.02	-.04	.00	.02	.02
	Background Noise → Excitement	-.01	.05	-.33	.03	.05	1.30
	Device Used[b] → Excitement	.04	.05	1.32	.04	.05	1.45

Notes: Estimates represent standardized path coefficients. Significant paths have been highlighted in orange to improve orientation. The darker the color, the higher the significance level. [a] dummy coded: "0" = female, "1" = male; [b] dummy coded: "0" = speaker, "1" = headphones; *p < .1; ** p < .05; ***p < .001; av. = average; S.E. = standard errors.

5.2. Interpretation of Brand Voice Personality Model

The BVP-Model was constructed based on a conducted SEM with the speaker's gender-specific effects, in which the perceptions of the three brand voice personality dimensions, sincerity, sensitivity, and excitement, were related to eleven acoustical measures (see Appendix O). The acoustical measures of interest in this study derived from previous findings in psychology and phonetics regarding correlations between personality traits. They can be described according to the four distinct soundwave dimensions: timing, amplitude, frequency, and voice quality (Hildebrand et al., 2020; Jurafsky & Martin, 2020).

In accordance with the BVP-Model, the following voice profiles are derived for each brand voice personality dimension. This provides an overview of the combination of acoustic measures that is most favorable for the respective personality perception in female and male voices. The voice profiles assist marketers and specialists in voice technology in transcribing a brand personality into a voice.

5.2.1. Sincerity in Brand Voice

Sincerity in brands is perceived when *females* speak fluently and loudly in brightness and show creakiness and breathiness (see **Table 13A**). This profile can be observed in the sincere voice of sample 18 ($M_{sincerity}$ = 3.8; see Appendix A). The speaker shows fluent speech with an above-average speaking rate of 3.4syl/s ($M_{speaking\_rate}$ = 3.08syl/s within the female stimuli set). Additionally, her voice reveals an intense brightness as she articulates the single words clearly. The speaker's creakiness is audible through her use of vocal fry in the first half of her speech. Vocal fry is a creaky vocal quality that occurs when speakers lower their pitch to the lowest register they are capable of producing, typically when finishing a sentence (Anderson et al., 2014). In addition, the speaker breathes audibly between sentences, which causes her to exhale at the beginning of each sentence, resulting in a breathy vocal quality.

The profile for *male voices* is very similar to the female voice profile (see **Table 13B**). Thus, a creaky male voice with brightness and fluent speech relates to a sincere brand personality. The voice of sample 89 is rated the highest on sincerity ($M_{sincerity}$ = 3.85) and is thus illustrative of a voice suitable for representing a sincere male brand (see Appendix A). The speaker's vocal creakiness is audible, and the speaker speaks with good articulatory clarity, following the desired brightness in a sincere voice. With a speaking rate of 3.8syl/s, the speaker talks with an above-average fluency ($M_{speaking\_rate}$ = 3.31syl/s within the male stimuli set).

Furthermore, voices are perceived as sincere when the listeners are female and the speaker's voice is male. This gender-specific effect on the perception of male voices can be explained by gender stereotypes resulting from socialization (Trouvain et al., 2021). Previous research suggests that sincere personality traits may be more commonly associated with males than females. This assumption is supported by studies showing male personalities being perceived in terms of physical and emotional power (Addington, 1968). In contrast, female personalities are often perceived in terms of social skills (Addington, 1968).

Table 13. Voice Profile of Brand Voice Personality Sincerity per Gender

	Voice Parameters	
	Listener's Perception	**Acoustical Measure (Metric)**
A. Female	fluent speech	increased speaking rate (syl/s)
	breathiness	increased h1-h2 (dB) decreased CPPS (dB)
	creakiness brightness	flat spectral slope
B. Male	fluent speech	increased speaking rate (syl/s)
	creakiness brightness	flat spectral slope flat spectral tilt

Notes: CPPS = smoothed cepstral peak prominence; syl = syllable; s = second; dB = Decibel.

5.2.2. Sensitivity in Brand Voice

In contrast to the voice profiles of sincerity, the voice profiles for the perception of sensitive brands demonstrate more notable differences between the genders. The *female voice* for sensitive brands can be described as fluent, slow, and monotone. Further, the voice is low-pitched and breathy but not rough (see **Table 14A**). Thus, all four voice dimensions are decisive for sensitivity perception in female voices. The profile indicates a stereotypical perception of a sensitive female speaking with a soft voice and a pleasant speech rate. The softness in voice is achieved through a breathy voice quality, characterized by low glottal tension and audible expiration noises as the vocal folds do not close completely. The speaker of voice sample 43 ($M_{sensitivity}$ = 3.8) demonstrates the sensitive brand voice with her audible breathiness (h1-h2 = 4.97dB), low-pitched voice (f0 mean = 211Hz), and slow and monotone speaking style (articulation rate = 4.3; intensity variability = 13.2; see Appendix A).

Voice qualities dominate the *male voice* profile for sensitivity perceptions. Therefore, a sensitive brand should be represented by a male voice that speaks brightly with breathiness and creakiness. Roughness in male voices decreases sensitivity perceptions and should be avoided (see **Table 14B**). An illustrative voice for a sensitive brand is sample 64, which is the highest-rated male voice on sensitivity within the vocal stimuli set ($M_{sensitivity}$ = 3.3; see Appendix A). The speaker demonstrates a breathy voice quality (h1-h2 = 1.24dB; CPPS 11.2dB) with vocal fry, i.e., creakiness, in parts.

The perception of sensitivity is favorably influenced when listeners are male or have less income for both gender voice profiles. In addition, female voices are associated with a sensitive personality when listeners are older. The tendency for males to perceive sensitivity through female voices can be attributed to the influence of gender stereotypes. Historically, women have been associated with social skills and a relatively emotional nature (Addington, 1968). This image is likely more anchored among older listeners, reflecting a classic male-female role distribution in which women are primarily responsible for childcare (social skills) while men are engaged in work (physical skills).

Table 14. Voice Profile of Brand Voice Personality Sensitivity per Gender

	Voice Parameters	
	Listener's Perception	**Acoustical Measure (Metric)**
A. Female	fluent speech	decreased average duration of silent pauses (s)
	slow speech	decreased articulation rate (syl/s)
	monotone intonation through loudness	decreased intensity variability (dB)
	low-pitched voice	decreased f0 mean (Hz)
	breathiness	increased h1-h2 (dB) decreased CPPS (dB)
	no roughness	decreased shimmer (%)
B. Male	breathiness	increased h1-h2 (dB) decreased CPPS (dB)
	creakiness brightness	flat spectral tilt
	no roughness	decreased shimmer (%)

Notes: CPPS = smoothed cepstral peak prominence; f0 = fundamental frequency; syl = syllable; s = second; Hz = Hertz; dB = Decibel.

5.2.3. Excitement in Brand Voice

As for the voice profiles of sincerity, female and male voices for the perception of exciting brands show substantial similarity (see **Table 15**). Exciting voices of both genders are breathy and creaky but not rough. Moreover, a less loud voice in terms of brightness and a less fluent speaking style with longer average silent pauses induce excitement perceptions.

At first view, a contradiction in fluency exists for female voices, as the results show that an increased speaking rate is associated with excitement perceptions, which indicates a fluent speaking style. In contrast, the favorable longer average silent pauses indicate a less fluent speaking style. Upon closer examination of the vocal stimuli, it becomes apparent that females rated high on excitement, take longer silent pauses while inhaling, and speak quickly between them, such as in voice sample 22 ($M_{excitement}$ = 2.6; see Appendix A). This combination results in the speaker being perceived as exciting. Research in American English supports this finding, as an increased speaking rate, and therefore increased fluency, is perceived as energetic in female voices in correlation and cue manipulation studies (Addington, 1968; Aronovitch, 1976).

Moreover, the results demonstrate that male, younger, and less educated listeners perceive female and male voices as exciting. These results not only point to gender differences in the perception of excitement through voice but also highlight generational differences. Further, it is worth noting that most of the voices that received high ratings for excitement are from younger speakers, such as sample 59 (23 years old), sample 35 (23 years old), and sample 65 (19 years old; see Appendix A). This suggests that younger speakers may naturally speak in a way perceived as exciting, particularly by listeners from the same age group. Therefore, it is recommended to use younger female or male voices for exciting brands, as excitement perceptions are likely to increase. A young target group is also more likely to perceive the excitement in these voices.

Table 15. Voice Profile of Brand Voice Personality Excitement per Gender

	Voice Parameters	
	Listener's Perception	**Acoustical Measure (Metric)**
A. Female	fluent speech through speaking rate	increased speaking rate (syl/s)
	hesitant/ less fluent speech through the use of long silent pauses	increased average duration of silent pauses (s)
	breathiness	increased h1-h2 (dB)
	creakiness brightness	flat spectral slope
	no roughness	decreased shimmer (%)
B. Male	hesitant/ less fluent speech	increased average duration of silent pauses (s)
	breathiness	decreased CPPS (dB)
	creakiness brightness	flat spectral slope
	no roughness	decreased shimmer (%)

Notes: CPPS = smoothed cepstral peak prominence; syl = syllable; s = second; dB = Decibel.

6. Implications

6.1. Managerial Implications

The evolution of voice AI offers companies and brands numerous opportunities to engage in a dialogue with their customers and improve their brand experience. For instance, speaking to customers with a brand voice, as is possible through smart speakers, creates a social and natural interaction between the brand and the customer (Reeves & Nass, 1996). The "Computers as Social Actors" (CASA) concept proposes that humans automatically ascribe personality, emotions, and social norms to computers, which in turn affects the user's behavior (Metze et al., 2011; Nass & Brave, 2005). Recent research on voice-based AI has shown that the CASA concept is also applied in conversations with voice assistants, as customers act as if they were talking to real people (Pitardi & Marriott, 2021). The voice plays a central role here, as the voice is mainly responsible for the perceived humanness of voice-based conversational agents (Novak & Hoffman, 2019; Schweitzer et al., 2019). In this regard, the use of human voices is preferable to that of artificial ones, as human voices are considered to achieve enhanced levels of effectiveness, greater attention, and better recall with less concentration (Atkinson et al., 2005; Rodero, 2017). But while artificial voices sounded like computers in the past, nowadays, they are remarkably realistic and natural, as recently demonstrated through the voices of ChatGPT and the Voice Engine by OpenAI (OpenAI, 2023, 2024). Thus, integrating voice marketing into the overall marketing and branding strategy with a suitable brand voice is recommended, as interactions with anthropomorphized brands have been shown to influence product evaluations positively (Aggarwal & McGill, 2007), enhance product likeability (Chandler & Schwarz, 2010), and boost brand recall, affection, and loyalty (Rauschnabel & Ahuvia, 2014).

When a brand voice is established as a new touchpoint, it is advisable to identify a voice that distinguishes the brand from others and conveys the appropriate personality to create consistent brand communication. Most companies utilizing voice apps on smart speakers typically use the default voice of the assistant, e.g., Alexa, Siri, or Google Home. In addition to the evidence that natural voices are preferred in human-computer interactions, the use of default synthetic voices raises the possibility that the brands may be communicating a personality that does not align with the one they usually have (Atkinson et al., 2005; Chiou et al., 2020; Seaborn et al., 2022; Xu, 2019). For example, Amazon describes Alexa's personality as approachable, efficient, trustworthy, and natural (Amazon, 2024a). Further, in a perceptual study of voice assistants' personality, Siri was rated high on extraversion and agreeableness, and Cortana, Microsoft's voice assistant, was rated high on neuroticism (Hacker, 2021). The developed BVP-Scale offers marketers guidance in identifying their brand's voice personality, which further helps them translate their personality into a suitable brand voice. By selecting or designing an appropriate brand voice, customers can perceive the desired brand personality and differentiate the brand voice from the default voices of familiar voice assistants or other voice-based technology.

The way in which voice assistants and other conversational agents are perceived, including their personality and the gender assigned to them, may prove to be disadvantageous to brands. Most voice assistants are designed as female, or their default voices are set to be female (UNESCO, 2019). Developers explained this gender bias in voice AI by referring to research demonstrating that female and male listeners prefer female voices (Stern, 2017). In contrast, studies indicate that the voice of the opposite gender is preferred and that female users of smart speakers modify their voice assistant's voice to male, if possible (UNESCO, 2019; Zuckerman & Driver, 1988). Nevertheless, despite the possibility of modifying the voice of a voice assistant to align with a male or gender-neutral persona, brands traditionally associated with masculinity face greater challenges in conveying their personalities due to the prevalence of gender bias in voice-based technology. Therefore, companies of male brands should design or choose their voice to avoid communicating with a customer through a default female voice.

Communication with the brand's gender voice is essential, as research on voice-product congruence evaluations has shown. In their experiments, Efthymiou et al. (2024) demonstrate that a masculine (feminine) voice is associated with masculine (feminine) products. Further, they show that a conversational agent with a masculine voice improves the advertising performance of a stereotypically masculine product. In contrast, a feminine voice improves the advertising performance of a stereotypically feminine product (Efthymiou et al., 2024). Therefore, marketers should identify their brand's gender in accordance with the characteristics of the brand personality and the product. As shown by this study's developed BVP-Model, the voice's gender determines the precise combination of voice parameters that will result in the perception of a specific brand personality. The most significant differences in voice profiles were detected in the perception of sensitive female and male brands.

Finally, the application areas for brand voices are numerous, and with new sophisticated technology, more possibilities will arise. Voice-based AI is not restricted to smartphones and smart speakers but is utilized at all possible contact points with customers, as shown by several examples from practice. For instance, the clothing retailer H&M offers an integrated voice AI mirror that interacts with customers by providing fashion recommendations or taking selfies (Baron, 2018). Mercedes-Benz has integrated its voice assistant "MBUX" in several models, which assists while driving and anticipates preferred settings in the car (Mercedes-Benz Group AG, 2022). Further, computer-generated voices are utilized in interactive voice response (IVR) systems, typically used in telephone customer services by educational, health, financial, or governmental institutions (Inam et al., 2017). Despite the rise of chatbots in customer services, IVR services are still highly used, as customers prefer to call for problem-solving requests (Agarwal et al., 2023). Companies communicating with their customers at different voice-based brand touchpoints should consider tailoring their voice to their target audience. As this study showed, the listener's gender impacts the perception of each brand's personality. Prevailing gender stereotypes can explain these gender-specific perceptions. Other listeners' socio-demographic information, such as age, education, and income, can also play an essential role in brand perceptions. Tailoring voices based on audience characteristics is recommended when audiences differ significantly. Ad targeting could then use different voices for different audiences to increase the possibility that the desired brand personality is perceived through brand voices (Iqbal et al., 2023).

6.2. Theoretical Implications

This study makes three essential contributions to vocal personality perception and voice marketing research. The developed BVP-Scale and BVP-Model, combined with the derived brand voice profiles for sincerity, sensitivity, and excitement, shed light on how brand personality can be translated into a voice, thereby supporting and expanding existing theories and concepts of brand personality perception.

This study is the first to examine the perception of brand personalities through voice. Literature on the vocal perception of personalities mainly focuses on the perception of human personalities (Charoenruk & Olson, 2018; Rosenberg & Hirschberg, 2009; Scherer, 1978). However, as human personalities can only be transferred to a limited extent to brands, a separate investigation of brand personality perception through voice is needed. The results of this study extend the knowledge on the perception of well-known brand personality dimensions and traits by showing that three brand personality dimensions exist, which can be perceived through voice alone: sincerity, sensitivity, and excitement. These brand voice personality dimensions can be measured with the developed BVP-Scale. The present study provides further evidence supporting the theory that a limited number of key personalities can be perceived through voice, which are combined in what are known as halo clusters (McAleer & Belin, 2019). These clusters represent a specific brand voice personality dimension.

Secondly, this study applies a multivariate analysis of voice parameters by examining the influence of linear combinations of voice parameters on brand personality perceptions. In natural speech, information is provided through the verbal channel, which encodes the semantic content of a message, and through the paraverbal channel, i.e., the acoustic parameters of the voice (Apple et al., 1979; Ketrow, 1990). This paralinguistic information is communicated through the combination and interplay of different voice parameters, which, on the one hand, makes every voice sound unique and, on the other hand, induces perceptions of the speaker's identity and personality (Belin et al., 2011; Rodero, 2013). The importance of investigating interaction effects of voice parameters on personality perception was already addressed as the "most difficult problem" (Addington, 1968, p. 503) and formulated as a future research recommendation in early phonetic studies, such as the study by Addington (1968) and Ray (1986). Also, Gobl and Chasaide (2003) addressed this issue in their research on the effect of voice quality on emotion, mood, and attitude. They suggested that "methodological limitations" (Gobl & Chasaide, 2003, p. 207) have been the reason for the lack or failure of multivariate analyses of voice parameters. This study is one of the few to react to the need to investigate the interaction effects of voice parameters on personality perception (Feinberg et al., 2005; McAleer et al., 2014; Pisanski & Rendall, 2011; Wu et al., 2021). The developed BVP-Model expands the knowledge of how combining specific voice parameters may influence personality judgments compared to when voice parameters are considered in isolation.

Finally, this study examines only objectively measurable voice parameters of all four soundwave dimensions of the voice (i.e., timing, frequency, amplitude, and voice quality), making the results more comparable between perceptual studies. Whereas the acoustic measures of the timing and frequency dimension, e.g., speaking rate and f0, were primarily utilized in perceptual studies, voice qualities, such as breathiness, creakiness, and hoarseness, were previously subjectively evaluated or neglected. (Hildebrand et al., 2020; Klasmeyer & Sendlmeier, 1997). However, the examination of voice qualities is essential in perceptual studies, as these parameters contribute significantly to the uniqueness of a speaker's voice (Pearsell & Pape, 2023). By incorporating acoustical measures of spectral features of the voice, such as HNR, CPPS, and h1-h2, this study follows the methodological approaches of recent research in vocal personality perception and thus expands the knowledge on voice qualities perceptions (Koutsoumpis & Vries, 2022; McAleer et al., 2014; Wu et al., 2021).

7. Limitations and Future Recommendations

This study employed a correlational study design and used external judgments from naïve listeners to investigate how brand personalities were perceived through voices developing the BVP-Scale and BVP-Model. The exploratory nature of this study leads to limitations that also provide opportunities for future research.

The 96 voices of the semi-spontaneous speech samples from the JESS represent one of the biggest limitations. For one thing, the JESS has a specific age distribution, as voices from speakers between 26-59 years were missing. The speech corpus is still appropriate for this perceptual study, as the perceived age of the speakers was documented, and middle-aged-sounding voices were covered. Moreover, voices showed sufficient variability in the 15 initial acoustical measures of interest, but extreme, highly pitched, or hoarse voices were still missing in the JESS. Since voices must be recorded under the same conditions to analyze and compare them phonetically, using speech corpora developed under controlled conditions is necessary. Although the linguistic community offers a variety of speech corpora, they often are not appropriate for voice research as they do not contain standardized voice samples (Zäske et al., 2020). It is recommended that future research focus on developing a speech corpus with the specific aim of investigating brand personality perception. This corpus should include a range of speakers of varying genders and ages and a variety of voices in terms of their acoustical measures.

Further, the speech content was controlled using neutral content, i.e., a description of a farm scene. While some studies argue that the content of speech influences the perception of personality (Apple et al., 1979; Tsantani et al., 2016), other studies argue against the decisive influence of content (Ketrow, 1990; Tusing & Dillard, 2000). Also, the context of personality perception may be necessary. For example, the study of Tigue et al. (2012) demonstrated that people were more sensitive to vocal cues of dominance during wartime than during times without war, influencing their voting behavior. In further studies, the content of speech and the situational context of the listeners should be investigated as possible factors affecting the perception of brand personalities through voices.

The present study focused on a particular language area, employing German voices and brand personality traits. To ascertain the cross-cultural validity of the findings, it is recommended that the study be replicated in English-speaking countries and countries that differ significantly in their communication models. It is anticipated that differences in brand personality perceptions will emerge across countries as individuals learn to utilize their voices in culturally determined ways (Krauss et al., 2002; Waaramaa et al., 2021). Insights into the perception of brand personality in other countries could inform decisions regarding the adaptation of a brand voice for international sales markets. This decision can have marketing implications in terms of potential synergy effects.

Essay 2 References

Aaker, J. L. (1997). Dimensions of Brand Personality. *Journal of Marketing Research*, 34(3), 347–356. https://doi.org/10.1177/002224379703400304

Aaker, J. L. (2019). Dimensionen der Markenpersönlichkeit. In *Handbuch Markenführung* (pp. 1199–1209). Springer Gabler, Wiesbaden. https://doi.org/10.1007/978-3-658-13342-9_63

Addington, D. W. (1968). The relationship of selected vocal characteristics to personality perception. Speech Monographs, 35(4), 492–503. https://doi.org/10.1080/03637756809375599

Agarwal, R., Kline, P., & Raabe, J. (2023). *Why IVR still matters in an AI world. McKinsey & Company*. https://www.mckinsey.com/capabilities/operations/our-insights/operations-blog/why-ivr-still-matters-in-an-ai-world

Aggarwal, P., & McGill, A. L. (2007). Is that car smiling at me? Schema congruity as a basis for evaluating anthropomorphized products. *Journal of Consumer Research*, 34(4), 468–479. https://doi.org/10.1086/518544

Allport, G. W., & Cantril, H. (1934). Judging personality from voice. *The Journal of Social Psychology*, 5(1), 37–55.

Amazon. (2022). *What Are Alexa Skills?* https://developer.amazon.com/en-US/alexa/alexa-skills-kit

Amazon. (2024). *The Alexa Personality*. https://developer.amazon.com/en-US/alexa/branding/alexa-guidelines/communication-guidelines/brand-voice

Anderson, R. C., Klofstad, C. A., Mayew, W. J., & Venkatachalam, M. (2014). Vocal fry may undermine the success of young women in the labor market. *PLOS ONES*, 9(5), e97506. https://doi.org/10.1371/journal.pone.0097506

Anjos de Oliveira, A., Dajer, M., Fernandes, P., & Teixeira, J. (2020). Clustering of Voice Pathologies based on Sustained Voice Parameters. In *13th International Joint Conference on Biomedical Engineering Systems and Technologies* (Vol. 4, pp. 280–287). https://doi.org/10.5220/0009146202800287

Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, 37(5), 715–727. https://doi.org/10.1037/0022-3514.37.5.715

Aronovitch, C. D. (1976). The voice of personality: Stereotyped judgments and their relation to voice quality and sex of speaker. *The Journal of Social Psychology*, 99(2), 207–220. https://doi.org/10.1080/00224545.1976.9924774

Atari, M., & Henrich, J. (2023). Historical Psychology. *Current Directions in Psychological Science*, 32(2), 176–183. https://doi.org/10.1177/09637214221149737

Atkinson, R. K., Mayer, R. E., & Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology*, 30(1), 117–139. https://doi.org/10.1016/j.cedpsych.2004.07.001

Azoulay, A., & Kapferer, J.-N. (2003). Do brand personality scales really measure brand personality? *Journal of Brand Management*, 11(2), 143–155. https://doi.org/10.1057/palgrave.bm.2540162

Backhaus, K., Erichson, B., Gensler, S., Weiber, R., & Weiber, T. (2021). *Multivariate Analysemethoden*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-32425-4

Backhaus, K., Erichson, B., & Weiber, R. (2015). *Fortgeschrittene Multivariate Analysemethoden.* Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-46087-0

Baker, C. P., Sundberg, J., Purdy, S. C., Rakena, T. O., & Leão, S. H. d. S. (2022). CPPS and Voice-Source Parameters: Objective Analysis of the Singing Voice. *Journal of Voice*, 38(3), 549–560. https://doi.org/10.1016/j.jvoice.2021.12.010

Baron, K. (2018, June 7). Ambient Tech That Actually Works: H&M Launches A Voice Activated Mirror. *Forbes*. https://www.forbes.com/sites/katiebaron/2018/06/07/ambient-tech-that-actually-works-hm-launches-a-voice-activated-mirror/

Barsties V Latoszek, B., Bodt de, M., Gerrits, E., & Maryn, Y. (2018). The Exploration of an Objective Model for Roughness With Several Acoustic Markers. Journal of Voice: Official *Journal of the Voice Foundation*, 32(2), 149–161. https://doi.org/10.1016/j.jvoice.2017.04.017

Barsties V Latoszek, B., Maryn, Y., Gerrits, E., & Bodt de, M. (2017). The Acoustic Breathiness Index (ABI): A Multivariate Acoustic Model for Breathiness. *Journal of Voice*, 31(4), 511.e11-511.e27. https://doi.org/10.1016/j.jvoice.2016.11.017

Barsties V Latoszek, B., Maryn, Y., Gerrits, E., & Bodt de, M. (2018). A Meta-Analysis: Acoustic Measurement of Roughness and Breathiness. *Journal of Speech, Language, and Hearing Research*, 61(2), 298–323. https://doi.org/10.1044/2017_JSLHR-S-16-0188

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Statistical Psychology*, 3(2), 77–85. https://doi.org/10.1111/j.2044-8317.1950.tb00285.x

Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding voice perception. *British Journal of Psychology*, 102(4), 711–725. https://doi.org/10.1111/j.2044-8295.2011.02041.x

Blasi, D. E., Henrich, J., Adamou, E., Kemmerer, D., & Majid, A. (2022). Over-reliance on English hinders cognitive science. *Trends in Cognitive Sciences*, 26(12), 1153–1170. https://doi.org/10.1016/j.tics.2022.09.015

Boersma, P., & Weenink, D. (1992-2022*). Praat: doing phonetics by computer* (Version 6.1.50) [Computer software]. https://www.praat.org

Brakus, J. J., Schmitt, B. H., & Zarantonello, L. (2009). Brand Experience: What is It? How is it Measured? Does it Affect Loyalty? *Journal of Marketing*, 73(3), 52–68. https://doi.org/10.1509/jmkg.73.3.052

Brown, B. L., Giles, H., & Thakerar, J. N. (1985). Speaker evaluations as a function of speech rate, accent and context. *Language & Communication*, 5(3), 207–220. https://doi.org/10.1016/0271-5309(85)90011-4

Brown, B. L., Strong, W. J., & Rencher, A. C. (1975). Acoustic determinants of perceptions of personality from speech. *Linguistics*, 13(166), 11–32. https://doi.org/10.1515/ling.1975.13.166.11

Bryant, G. A. (2022). Vocal communication across cultures: Theoretical and methodological issues. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 377(1841), 20200387. https://doi.org/10.1098/rstb.2020.0387

Caprara, G. V., Barbaranelli, C., & Guido, G. (2001). Brand personality: How to make the metaphor fit? *Journal of Economic Psychology*, 22(3), 377–395. https://doi.org/10.1016/s0167-4870(01)00039-3

Carpenter, S. (2018). Ten Steps in Scale Development and Reporting: A Guide for Researchers. *Communication Methods and Measures*, 12(1), 25–44. https://doi.org/10.1080/19312458.2017.1396583

Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, 1(2), 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Chandler, J., & Schwarz, N. (2010). Use does not wear ragged the fabric of friendship: Thinking of objects as alive makes people less willing to replace them. *Journal of Consumer Psychology*, 20(2), 138–145. https://doi.org/10.1016/j.jcps.2009.12.008

Charoenruk, N., & Olson, K. (2018). Do Listeners Perceive Interviewers' Attributes from their Voices and Do Perceptions Differ by Question Type? *Field Methods*, 30(4), 312–328. https://doi.org/10.1177/1525822X18784500

Chiou, E. K., Schroeder, N. L., & Craig, S. D. (2020). How we trust, perceive, and learn from virtual humans: The influence of voice quality. *Computers & Education*, 146, 103756. https://doi.org/10.1016/j.compedu.2019.103756

Clark, J., Yallop, C., & Fletcher, J. (2007). *An introduction to phonetics and phonology* (3rd ed.). Wiley-Blackwell.

Conrad, F. G., Schober, M. F., & Dijkstra, W. (2008). Cues of Communication Difficulty in Telephone Interviews. In Lepkowski, J. M., Tucker, C., J. M. Brick, E. de Leeuw, L. Japec, P. J. Lavrakes, M. W. L. Link, & R. L. Sangster (Eds.), *Advances in Telephone Survey Methodology* (pp. 212–230). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470173404.ch10

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. https://doi.org/10.1007/BF02310555

Dahl, D. W. (2010). Understanding the role of spokesperson voice in broadcast advertising. In A. Krishna (Ed.), *Sensory marketing: Research on the sensuality of products* (pp. 169–182). Routledge. https://doi.org/10.4324/9780203892060-21

Dejonckere, P. H., Obbens, C., Moor, G. M. de, & Wieneke, G. H. (1993). Perceptual evaluation of dysphonia: Reliability and relevance. *Folia Phoniatrica*, 45(2), 76–83. https://doi.org/10.1159/000266220

DeVellis, R. F. (2017). *Scale Development: Theory and Applications*. (4th ed.). SAGE Publications.

Duvvuru, S., & Erickson, M. (2013). The effect of change in spectral slope and formant frequencies on the perception of loudness. *Journal of Voice: Official Journal of the Voice Foundation*, 27(6), 691–697. https://doi.org/10.1016/j.jvoice.2013.05.004

Dziuban, C. D., & Shirkey, E. C. (1974). When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin*, 81(6), 358–361. https://doi.org/10.1037/h0036316

Efthymiou, F., Hildebrand, C., Bellis, E. de, & Hampton, W. H. (2024). The Power of AI-Generated Voices: How Digital Vocal Tract Length Shapes Product Congruency and Ad Performance. *Journal of Interactive Marketing*, 59(2), 117–134. https://doi.org/10.1177/10949968231194905

Farrus, M., Hernando, J., & Ejarque, P. (2007). Jitter and Shimmer Measurements for Speaker Recognition. In *Interspeech 2007* (pp. 778–781). https://doi.org/10.21437/interspeech.2007-147

Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, 69(3), 561–568. https://doi.org/10.1016/j.anbehav.2004.06.012

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research,* 18(1), 39–50. https://doi.org/10.1177/002224378101800104

Fraile, R., & Godino-Llorente, J. I. (2014). Cepstral peak prominence: A comprehensive analysis. *Biomedical Signal Processing and Control*, 14, 42–54. https://doi.org/10.1016/j.bspc.2014.07.001

Frühholz, S., & Belin, P. (2019). The Science of Voice Perception. In S. Frühholz & P. Belin (Eds.), *The Oxford Handbook of Voice Perception* (1st ed., pp. 2–14). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198743187.013.1

Geuens, M., Weijters, B., & Wulf, K. de (2009). A new measure of brand personality. *International Journal of Research in Marketing*, 26(2), 97–107. https://doi.org/10.1016/j.ijresmar.2008.12.002

Gobl, C., & Chasaide, A. N. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, 40(1-2), 189–212. https://doi.org/10.1016/S0167-6393(02)00082-1

Gordon, J. K., Andersen, K., Perez, G., & Finnegan, E. (2019). How Old Do You Think I Am? Speech-Language Predictors of Perceived Age and Communicative Competence. *Journal of Speech, Language, and Hearing Research: JSLHR*, 62(7), 2455–2472. https://doi.org/10.1044/2019_JSLHR-L-19-0025

Gordon, M., & Ladefoged, P. (2001). Phonation types: a cross-linguistic overview. *Journal of Phonetics*, 29(4), 383–406. https://doi.org/10.1006/jpho.2001.0147

Grohmann, B. (2009). Gender Dimensions of Brand Personality. Journal of Marketing Research, 46(1), 105–119. https://doi.org/10.1509/jmkr.46.1.105

Guha, A., Bressgott, T., Grewal, D., Mahr, D., Wetzels, M., & Schweiger, E. (2023). How artificiality and intelligence affect voice assistant evaluations. *Journal of the Academy of Marketing Science*, 51(4), 843–866. https://doi.org/10.1007/s11747-022-00874-7

Hacker, A. (2021). "Alexa, who are you?" – Analysing Alexa's, Cortana's and Siri's Vocal Personality. In S. Hillmann, B. Weiss, T. Michael, & S. Möller (Eds.), *Konferenz Elektronische Sprachsignalverarbeitung* (pp. 11–18).

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2013). *Multivariate data analysis* (7th ed.). Pearson Education Limited.

Hair, J. F., Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, 109, 101–110. https://doi.org/10.1016/j.jbusres.2019.11.069

Hardesty, L. (2019). The history of Amazon's recommendation algorithm. Collaborative filtering and beyond. *Amazon Science*. https://www.amazon.science/the-history-of-amazons-recommendation-algorithm

Harnsberger, J. D., Shrivastav, R., Brown, W. S., Rothman, H., & Hollien, H. (2008). Speaking rate and fundamental frequency as speech cues to perceived age. *Journal of Voice*, 22(1), 58–69. https://doi.org/10.1016/j.jvoice.2006.07.004

Hildebrand, C., Efthymiou, F., Busquet, F., Hampton, W. H., Hoffman, D. L., & Novak, T. P. (2020). Voice analytics in business research: Conceptual foundations, acoustic feature extraction, and applications. *Journal of Business Research*, 121, 364–374. https://doi.org/10.1016/j.jbusres.2020.09.020

Hillenbrand, J., & Houde, R. A. (1996). Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *Journal of Speech and Hearing Research*, 39(2), 311–321. https://doi.org/10.1044/jshr.3902.311

Hodges-Simeon, C. R., Gaulin, S. J. C., & Puts, D. A. (2010). Different Vocal Parameters Predict Perceptions of Dominance and Attractiveness. *Human Nature*, 21(4), 406–427. https://doi.org/10.1007/s12110-010-9101-5

Hogan, R., & Bond, M. H. (2009). Culture and personality. In P. Corr & G. Matthews (Eds.), *The Cambridge Handbook of Personality Psychology* (pp. 577–588). Cambridge: Cambridge University Press.

Homburg, C., Schwemmle, M., & Kuehnl, C. (2015). New Product Design: Concept, Measurement, and Consequences. *Journal of Marketing*, 79(3), 41–56. https://doi.org/10.1509/jm.14.0199

Hörner, T. (2023). Marken auf Sprachassistenten. In T. Hörner (Ed.), *Marketing mit Sprachassistenten: So setzen Sie Alexa & Co strategisch erfolgreich ein* (2nd ed., pp. 191–222). Springer Gabler. https://doi.org/10.1007/978-3-658-40196-2_6

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and Deterring Insufficient Effort Responding to Surveys. *Journal of Business and Psychology*, 27(1), 99–114. https://doi.org/10.1007/s10869-011-9231-8

Inam, I. A., Azeta, A. A., & Daramola, O. (2017). Comparative analysis and review of interactive voice response systems. In *2017 Conference on Information Communication Technology and Society (ICTAS)* (pp. 1–6). IEEE. https://doi.org/10.1109/ictas.2017.7920660

Iqbal, U., Bahrami, P. N., Trimananda, R., Cui, H., Gamero-Garrido, A., Dubois, D. J., Choffnes, D., Markopoulou, A., Roesner, F., & Shafiq, Z. (2023). Tracking, Profiling, and Ad Targeting in the Alexa Echo Smart Speaker Ecosystem. In *Proceedings of the 2023 ACM on Internet Measurement Conference* (pp. 569–583). https://doi.org/10.1145/3618257.3624803

Jandura, O. (2018). Fake Data? In Rössler, P., & Rossmann, C. (Ed.), *Kumulierte Evidenzen* (pp. 207–223). Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-18859-7_10

Johnstone, t., & Scherer, K. R. (2000). Vocal communication of emotion. In Lewis M. & Haviland J. (Eds.), *Handbook of Emotion* (2nd ed.). New York: Guilford.

Jurafsky, D., & Martin, J. H. (2020). *Speech and language processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (3rd ed.).

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. https://doi.org/10.1007/BF02291575

Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, 34(1), 111–117. https://doi.org/10.1177/001316447403400115

Ketrow, S. M. (1990). Attributes of a telemarketer's voice and persuasiveness.: A review and synthesis of the literature. *Journal of Direct Marketing*, 4(3), 7–21. https://doi.org/10.1002/dir.4000040304

Klasmeyer, G., & Sendlmeier, W. F. (1997). The classification of different phonation types in emotional and neutral speech. *International Journal of Speech, Language and the Law*, 4(1), 104–124. https://doi.org/10.1558/ijsll.v4i1.104

Koutsoumpis, A., & Vries, R. E. de (2022). What Does Your Voice Reveal About You? Trait Activation of Voice Characteristics and Their Relation With Personality and Communication Styles. *Journal of Individual Differences*, 43(3), 160–167. https://doi.org/10.1027/1614-0001/a000362

Kramer, E. (1964). Personality Stereotypes in Voice: A Reconsideration of the Data. *The Journal of Social Psychology*, 62(2), 247–251. https://doi.org/10.1080/00224545.1964.9919523

Krauss, R. M., Freyberg, R., & Morsella, E. (2002). Inferring speakers' physical attributes from their voices. *Journal of Experimental Social Psychology*, 38(6), 618–625. https://doi.org/10.1016/S0022-1031(02)00510-3

Kreiman, J., & Sidtis, D. (2011). Perception of Emotion and Personality from Voice. In *Foundations of Voice Studies* (pp. 302–360). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444395068.ch9

Kuang, J., & Liberman, M. (2018). Integrating Voice Quality Cues in the Pitch Perception of Speech and Non-speech Utterances. *Frontiers in Psychology*, 9, 2147. https://doi.org/10.3389/fpsyg.2018.02147

Kumar, A. (2018). Story of Aaker's brand personality scale criticism. *Spanish Journal of Marketing - ESIC*, 22(2), 203–230. https://doi.org/10.1108/SJME-03-2018-005

Kumar, V., Dixit, A., Javalgi, R. G., & Dass, M. (2016). Research framework, strategies, and applications of intelligent agent technologies (IATs) in marketing. *Journal of the Academy of Marketing Science*, 44(1), 24–45. https://doi.org/10.1007/s11747-015-0426-9

Lee, H., & Cho, C.-H. (2020). Uses and gratifications of smart speakers: Modelling the effectiveness of smart speaker advertising. *International Journal of Advertising*, 39(7), 1150–1171. https://doi.org/10.1080/02650487.2020.1765657

Leiner, D. J. (2019). Too Fast, too Straight, too Weird: Non-Reactive Indicators for Meaningless Data in Internet Surveys. *Survey Research Methods*, 13(3), 229–248. https://doi.org/10.18148/srm/2019.v13i3.7403

Lieven, T. (2014). Vom Markengeschlecht zur Markenstärke. *Marketing Review St. Gallen*, 31(4), 80–88. https://doi.org/10.1365/s11621-014-0384-5

Lorenzo-Seva, U. (2021). Solomon: A method for splitting a sample into equivalent subsamples in factor analysis. *Behavior Research Methods*, 54(6), 2665–2677. https://doi.org/10.3758/s13428-021-01750-y

Mayer, J. (2019-2021). *Praat-Stimmprofil: v2.3.0* [Praat plugin].

McAleer, P., & Belin, P. (2019). The Perception of Personality Traits from Voices. In S. Frühholz & P. Belin (Eds.), *The Oxford Handbook of Voice Perception* (1st ed., pp. 585–606). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198743187.013.26

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *PLOS ONE*, 9(3), 1–9. https://doi.org/10.1371/journal.pone.0090779.t001

Mehrabian, A., & Ferris, S. R. (1967). Inference of attitudes from nonverbal communication in two channels. *Journal of Consulting Psychology*, 31(3), 248–252. https://doi.org/10.1037/h0024648

Melzner, J., Bonezzi, A., & Meyvis, T. (2022). EXPRESS: Information Disclosure in the Era of Voice Technology. *Journal of Marketing*, 87(4), 491-509. https://doi.org/10.1177/00222429221138286

Mercedes-Benz Group AG. (2022, April 19). *MBUX reaches a new level* [Press release]. https://group.mercedes-benz.com/innovation/digitalisation/connectivity/mbux-interior-assist.html

Mertens, P. (2022). The Prosogram Model for Pitch Stylization and Its Applications in Intonation Transcription. In J. A. Barnes & S. Shattuck-Hufnagel (Eds.), *Prosodic Theory and Practice.* (pp. 259–286). MA: MIT Press. https://doi.org/10.7551/mitpress/10413.003.0010

Metze, F., Black, A., & Polzehl, T. (2011). A Review of Personality in Voice-Based Man Machine Interaction. In J. Jacko (Ed.), *Human-Computer Interaction. Interaction Techniques and Environments - 14th International Conference, HCI International 2011* (pp. 358–367). https://doi.org/10.1007/978-3-642-21605-3_40

Moore, W. E. (1939). Personality Traits and Voice Quality Deficiencies. *Journal of Speech Disorders*, 4(1), 33–36. https://doi.org/10.1044/jshd.0401.33

Morgado, F. F. R., Meireles, J. F. F., Neves, C. M., Amaral, A. C. S., & Ferreira, M. E. C. (2017). Scale development: Ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexão E Crítica*, 30(3). https://doi.org/10.1186/s41155-016-0057-1

Nass, C., & Brave, S. (2005). *Wired for speech: How voice activates and advances the human-computer relationship*. Cambridge: MIT Press.

Novak, T. P., & Hoffman, D. L. (2019). Relationship journeys in the internet of things: A new framework for understanding interactions between consumers and smart objects.

Journal of the Academy of Marketing Science, 47(2), 216–237. https://doi.org/10.1007/s11747-018-0608-3

Oksenberg, L., Coleman, L., & Cannell, C. F. (1986). Interviewers' Voices and Refusal Rates in Telephone Surveys. *Public Opinion Quarterly*, 50(1), 97. https://doi.org/10.1086/268962

Oleszkiewicz, A., Pisanski, K., Lachowicz-Tabaczek, K., & Sorokowska, A. (2017). Voice-based assessments of trustworthiness, competence, and warmth in blind and sighted adults. *Psychonomic Bulletin & Review*, 24(3), 856–862. https://doi.org/10.3758/s13423-016-1146-y

OpenAI. (2023). ChatGPT can now see, hear, and speak. *OpenAI.* https://openai.com/index/chatgpt-can-now-see-hear-and-speak/

OpenAI. (2024). Navigating the challenges and opportunities of synthetic voices. *OpenAI.* https://openai.com/index/navigating-the-challenges-and-opportunities-of-synthetic-voices/

Packard, G., & Berger, J. (2024). The Emergence and Evolution of Consumer Language Research. *Journal of Consumer Research*, 51(1), 42–51. https://doi.org/10.1093/jcr/ucad013

Pear, T. H. (1931). *Voice and Personality* (1st ed.). Chapman and Hall, Ltd. https://doi.org/10.1038/128657a0

Pearsell, S., & Pape, D. (2023). The effects of different voice qualities on the perceived personality of a speaker. *Frontiers in Communication*, 7. https://doi.org/10.3389/fcomm.2022.909427

Peng, Y., Zebrowitz, L. A., & Lee, H. K. (1993). The Impact of Cultural Background and Cross-Cultural Experience on Impressions of American and Korean Male Speakers. *Journal of Cross-Cultural Psychology*, 24(2), 203–220. https://doi.org/10.1177/0022022193242005

Peterson, R. A., Cannito, M. P., & Brown, S. P. (1995). An Exploratory Investigation of Voice Characteristics and Selling Effectiveness. *Journal of Personal Selling & Sales Management*, 15(1), 1–15. https://doi.org/10.1080/08853134.1995.10754008

Pisanski, K., & Rendall, D. (2011). The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size, masculinity, and attractiveness. *The Journal of the Acoustical Society of America*, 129(4), 2201–2212. https://doi.org/10.1121/1.3552866

Pitardi, V., & Marriott, H. R. (2021). Alexa, she's not human but… Unveiling the drivers of consumers' trust in voice-based artificial intelligence. *Psychology & Marketing*, 38(4), 626–642. https://doi.org/10.1002/mar.21457

Pompino-Marschall, B. (2009). *Einführung in die Phonetik* (3rd ed.). Walter de Gruyter https://doi.org/10.1515/9783110224818

Rauschnabel, P. A., & Ahuvia, A. C. (2014). You're so lovable: Anthropomorphism and brand love. *Journal of Brand Management*, 21, 372–395. https://doi.org/10.1057/bm.2014.14

Ray, G. B. (1986). Vocally cued personality prototypes: An implicit personality theory approach. *Communication Monographs*, 53(3), 266–276. https://doi.org/10.1080/03637758609376141

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people*. Cambridge University Press.

Revelle, W. (2024). *psych: Procedures for Psychological, Psychometric, and Personality Research* (Version 2.4.3) [Computer software]. https://CRAN.R-project.org/package=psych

Riding, D., Lonsdale, D., & Brown, B. (2006). The Effects of Average Fundamental Frequency and Variance of Fundamental Frequency on Male Vocal Attractiveness to Women. *Journal of Nonverbal Behavior*, 30(2), 55–61. https://doi.org/10.1007/s10919-006-0005-3

Rodero, E. (2012). A comparative analysis of speech rate and perception in radio bulletins. *Text & Talk*, 32(3), 391–411. https://doi.org/10.1515/text-2012-0019

Rodero, E. (2013). The Perception of a Broadcasting Voice. *US-China Education Review*, 3(4), 225–230.

Rodero, E. (2017). Effectiveness, attention, and recall of human and artificial voices in an advertising story. Prosody influence and functions of voices. *Computers in Human Behavior,* 77, 336–346. https://doi.org/10.1016/j.chb.2017.08.044

Rosenberg, A., & Hirschberg, J. (2009). Charisma perception from text and speech. *Speech Communication*, 51(7), 640–655. https://doi.org/10.1016/j.specom.2008.11.001

Rosseel, Y. (2012). Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software,* 48(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Sarstedt, M., & Mooi, E. (2019). Regression Analysis. In *A Concise Guide to Market Research* (pp. 209–256). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-56707-4_7

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2020). Partial Least Squares Structural Equation Modeling. In C. Homburg, M. Klarmann, & A. Vomberg (Eds.), *Springer eBook Collection Business and Management. Handbook of market research* (pp. 1–47). Springer. https://doi.org/10.1007/978-3-319-05542-8_15-2

Scherer, K. R. (1972). Judging personality from voice: A cross-cultural approach to an old issue in interpersonal perception. *Journal of Personality*, 40(2), 191–210. https://doi.org/10.1111/j.1467-6494.1972.tb00998.x

Scherer, K. R. (1974). Voice quality analysis of American and German speakers. *Journal of Psycholinguistic Research,* 3(3), 281–298. https://doi.org/10.1007/BF01069244

Scherer, K. R. (1978). Personality inference from voice quality: The loud voice of extroversion. *European Journal of Social Psychology*, 8(4), 467–487. https://doi.org/10.1002/ejsp.2420080405

Schweinberger, S. R., & Zäske, R. (2019). Perceiving Speaker Identity from the Voice. In S. Frühholz & P. Belin (Eds.), *The Oxford Handbook of Voice Perception* (1st ed., pp. 538–560). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198743187.013.24

Schweitzer, F., Belk, R., Jordan, W., & Ortner, M. (2019). Servant, friend or master? The relationships users build with voice-controlled smart devices. *Journal of Marketing Management*, 35(7-8), 693–715. https://doi.org/10.1080/0267257X.2019.1596970

Seaborn, K., Miyake, N. P., Pennefather, P., & Otake-Matsuura, M. (2022). Voice in Human–Agent Interaction. *ACM Computing Surveys*, 54(4), 1–43. https://doi.org/10.1145/3386867

Shipp, T., Qi, Y., Huntley, R., & Hollien, H. (1992). Acoustic and temporal correlates of perceived age. *Journal of Voice*, 6(3), 211–216. https://doi.org/10.1016/S0892-1997(05)80145-6

Skoog Waller, S., & Eriksson, M. (2016). Vocal Age Disguise: The Role of Fundamental Frequency and Speech Rate and Its Perceived Effects. *Frontiers in Psychology*, 7. https://doi.org/10.3389/fpsyg.2016.01814

Sluijter, A. M., & van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America*, 100(4), 2471–2485. https://doi.org/10.1121/1.417955

Statista Market Insights. (2023). *Speakers*. https://www.statista.com/outlook/cmo/consumer-electronics/tv-radio-multimedia/speakers/worldwide#revenue

Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing Measurement Invariance in Cross-National Consumer Research. *Journal of Consumer Research*, 25(1), 78–107. https://doi.org/10.1086/209528

Stern, J. (2017, February 21). Alexa, Siri, Cortana: The Problem With All-Female Digital Assistants. *The Wallstreet Journal*. https://www.wsj.com/articles/alexa-siri-cortana-the-problem-with-all-female-digital-assistants-1487709068

Stévens, J. C., & Hall, J. W. (1966). Brightness and loudness as functions of stimulus duration. *Perception & Psychophysics*, 1(9), 319–327. https://doi.org/10.3758/BF03215796

Street, R. L., & Brady, R. M. (1982). Speech rate acceptance ranges as a function of evaluative domain, listener speech rate, and communication context. *Communication Monographs*, 49(4), 290–308. https://doi.org/10.1080/03637758209376091

Svec, J., Schutte H. K., & Šram F. (1999). Variability of vibration of normal vocal folds as seen in videokymography. In Dejonckere P.H. & Peters H.F.M. (Eds.), *Communication and Its Disorders: A Science In Progress. Proceedings 24th Congress International Association of Logopedics and Phoniatrics* (1st ed., pp. 122–125).

Tian, K. T., Bearden, W. O., & Hunter, G. L. (2001). Consumers' Need for Uniqueness: Scale Development and Validation. *Journal of Consumer Research*, 28(1), 50–66. https://doi.org/10.1086/321947

Tigue, C. C., Borak, D. J., O'Connor, J. J., Schandl, C., & Feinberg, D. R. (2012). Voice pitch influences voting behavior. *Evolution and Human Behavior*, 33(3), 210–216. https://doi.org/10.1016/j.evolhumbehav.2011.09.004

Trouvain, J., Weiss, B., & Barkat-Defradas, M. (2021). Voice Attractiveness: Concepts, Methods, and Data. In *Voice attractiveness: Studies on sexy, likable, and charismatic speakers* (pp. 3–16). Springer, Singapore. https://doi.org/10.1007/978-981-15-6627-1_1

Tusing, K. J., & Dillard, J. P. (2000). The sounds of dominance. *Human Communication Research*, 26(1), 148–171. https://doi.org/10.1111/j.1468-2958.2000.tb00754.x

UNESCO. (2019). I'd blush if I could: closing gender divides in digital skills through education. *United Nations Educational, Scientific and Cultural Organization*. https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=7

Waaramaa, T., Lukkarila, P., Järvinen, K., Geneid, A., & Laukkanen, A.-M. (2021). Impressions of Personality from Intentional Voice Quality in Arabic-Speaking and Native Finnish-Speaking Listeners. *Journal of Voice: Official Journal of the Voice Foundation*, 35(2), 326.e21-326.e28. https://doi.org/10.1016/j.jvoice.2019.09.003

Wagner, K., & Schramm-Klein, H. (2019). Alexa, Are You Human? Investigating Anthropomorphism of Digital Voice Assistants - A Qualitative Approach. In *Proceedings of the Fortieth International Conference on Information Systems (ICIS), Munich 2019.*

Wang, L., Huang, N., Hong, Y., Liu, L., Guo, X., & Chen, G. (2023). Voice-based AI in call center customer service: A natural field experiment. *Production and Operations Management*, 32(4), 1002–1018. https://doi.org/10.1111/poms.13953

Weiber, R., & Mühlhaus, D. (2014). *Strukturgleichungsmodellierung*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-35012-2

Wendahl, R. W. (1963). Laryngeal Analog Synthesis of Harsh Voice Quality. *Folia Phoniatrica,* 15(4), 241–250. https://doi.org/10.1159/000262968

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). Sage Publications, Inc.

Westlund, A. H., Källström, M., & Parmler, J. (2008). SEM-based customer satisfaction measurement: On multicollinearity and robust PLS estimation. *Total Quality Management & Business Excellence*, 19(7-8), 855–869. https://doi.org/10.1080/14783360802159527

Wiener, H. J. D., & Chartrand, T. L. (2014). The Effect of Voice Quality on Ad Efficacy. *Psychology & Marketing*, 31(7), 509–517. https://doi.org/10.1002/mar.20712

Wu, Q., Liu, Y., Li, D., Leng, H., Iqbal, Z., & Jiang, Z. (2021). Understanding one's character through the voice: Dimensions of personality perception from Chinese greeting word "Ni Hao". *The Journal of Social Psychology*, 161(6), 653–663. https://doi.org/10.1080/00224545.2020.1856026

Xu, K. (2019). First encounter with robot Alpha: How individual differences interact with vocal and kinetic cues in users' social responses. *New Media & Society*, 21(11-12), 2522–2547. https://doi.org/10.1177/1461444819851479

Zäske, R., Skuk, V. G., Golle, J., & Schweinberger, S. R. (2019). *The Jena Speaker Set (JESS) – A database of voice stimuli from unfamiliar young and old adult speakers: Supplementary Material*. https://osf.io/u6amw

Zäske, R., Skuk, V. G., Golle, J., & Schweinberger, S. R. (2020). The Jena Speaker Set (JESS)-A database of voice stimuli from unfamiliar young and old adult speakers. *Behavior Research Methods,* 52(3), 990–1007. https://doi.org/10.3758/s13428-019-01296-0

Zoghaib, A. (2017). The contribution of a brand spokesperson's voice to consumer-based brand equity. *Journal of Product & Brand Management*, 26(5), 492–502. https://doi.org/10.1108/JPBM-06-2016-1230

Zuckerman, M., & Driver, R. E. (1988). What sounds beautiful is good: The vocal attractiveness stereotype. *Journal of Nonverbal Behavior*, 13(2), 67–82. https://doi.org/10.1007/bf00990791

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99(3), 432–442. https://doi.org/10.1037//0033-2909.99.3.432

Essay 2 Appendix

Appendix A. Access to Vocal Stimulus Set of this Study

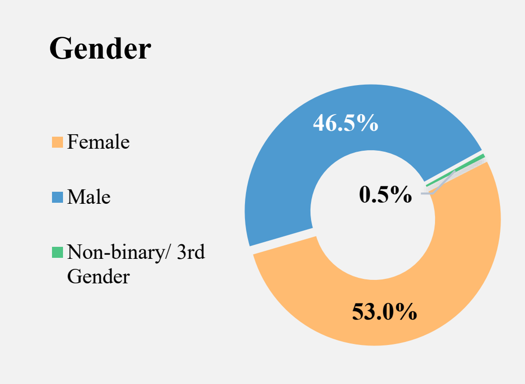
The voices used as this study's vocal stimuli were obtained from the Jena Speaker Set (JESS; Zäske et al., 2020). The vocal stimulus set is provided in the electronic appendix of this dissertation – in the folder named "Essay_2_Voice_Samples_Jena_Speaker_Set_JESS". The folder contains 98 audio files:

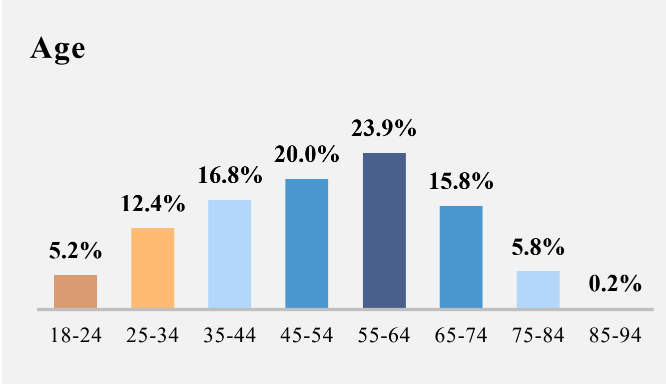
- The 2 audio files named "Skeine_fAN25" and "Skeine_mHD68" were used as the visual representation of the soundwaves of utterances of a female and male speaker to provide a better understanding of the voice dimensions and parameters described in section 2.2.1.
- The remaining 96 audio files represent the study's vocal stimulus set used for the online survey described in section 3.1. This study only used snippets of the semi-spontaneous speech recordings of the JESS describing a farmyard scene. Moreover, out of the 120 voices of the corpus, only 96 were used, as 24 voices exhibited audible dialects, speech disorders, or articulation difficulties. The corresponding files are each named according to the declared **voice sample number for this study** and the sample's *original file name from the JESS*, e.g., "**voicesample1_***fAB75*".

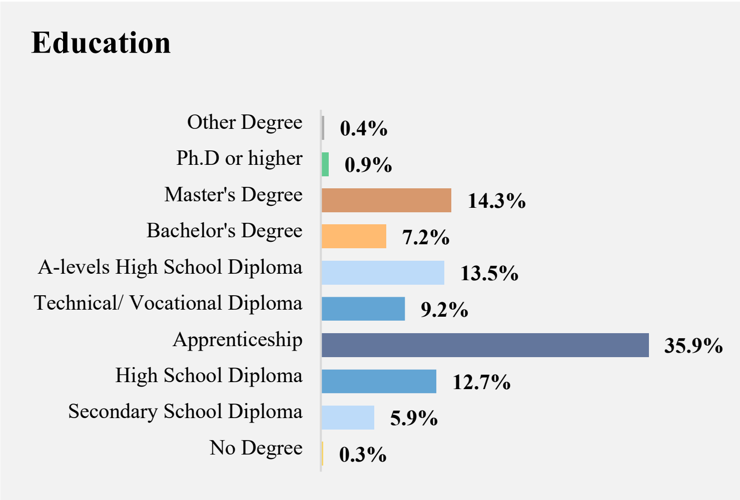
Appendix B. Summary of Pre-Test

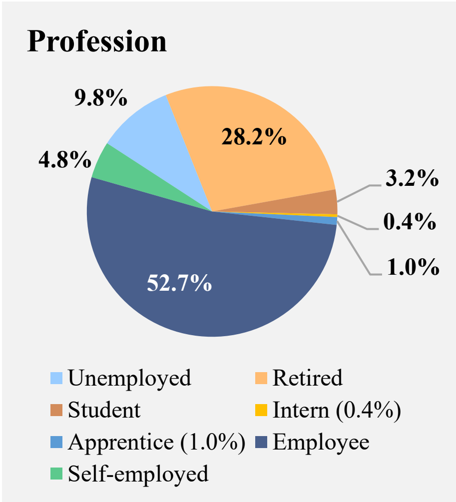
Purpose	Identifying the optimal number of voice ratings a participant could provide without losing focus, ensuring the highest quality of responses.
Method	Online questionnaire distributed through SurveyCircle
Procedure	Participants were asked to listen to and rate four voices, on average 27.5 seconds long, on a five-point scale (from 1 = ”does not apply at all” to 5 = ”applies completely”) on the extent to which they associated the voice with the 64 brand personality traits provided. Following each voice rating, participants were asked to rate their attention, concentration, and willingness to perform another rating using a five-point smiley face rating scale (1 = “very unhappy” to 5 = “very happy”). Participants who declined to rate another voice were directed to the final questions, which asked for general feedback at the end of the survey.
Sample Size	33 German participants (all participants rated at least two voices, 23 participants rated three voices, 19 participants rated all four voices)
Results	The data analysis on the attention, concentration, and willingness for further voice evaluations revealed that participants showed the lowest values after two voice ratings ($M_{attention} = 3.91$; $M_{concentration} = 3.76$; $M_{willingness} = 3.09$). The highest dropout rate of 10 participants was after two voice ratings. In addition, most participants indicated that a minimum of two (48.48%) and a maximum of three (36.36%) voice ratings would be appropriate. In conclusion, participants of the main study should rate two voices to obtain the highest quality responses.

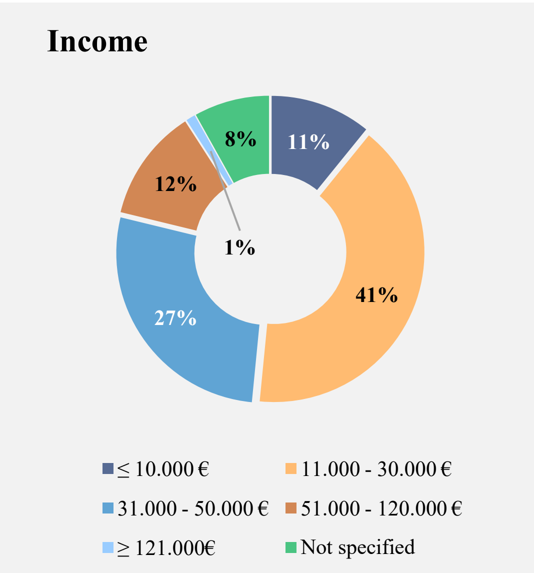
Appendix C. Sample Characteristics: Socio-Demographic & Survey-Related Data

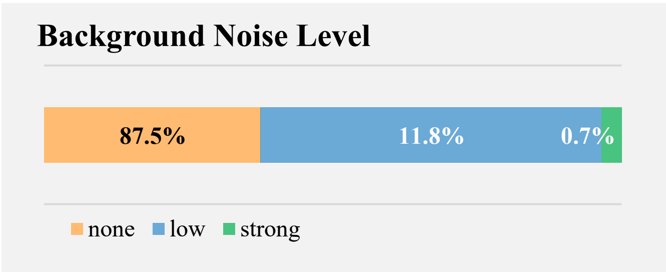


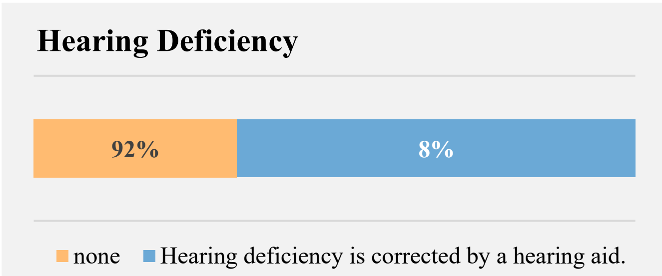


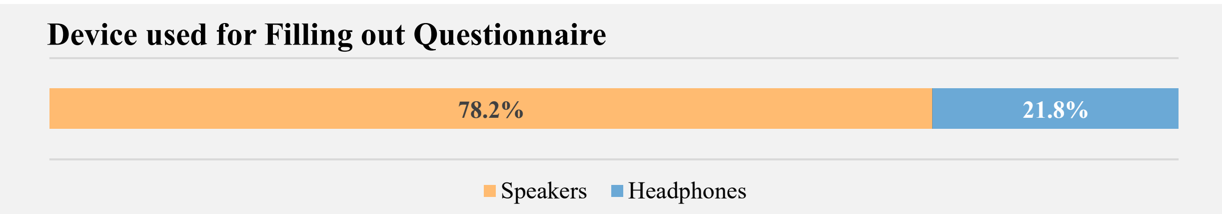












Notes: N = 2,000 participants.

Appendix D. Overview of Selected & Excluded Voices from the Jena Speaker Set

Female Speaker ID	Selected or Excluded with Explanation	Male Speaker ID	Selected or Excluded with Explanation
fAH18	Selected	mDA18	Selected
fBS18	Selected	mGR18	Selected
fES18	Selected	mJK18	Selected
fRB18	Selected	mRM18	Selected
fSL18	Excluded: stutter	mFS19	Selected
fAH19	Selected	mJK19	Selected
fJR20	Selected	mJB20	Selected
fLK20	Selected	mKK21	Selected
fJJ21	Excluded: interruptions due to laughter	mMB21	Selected
fKM21	Selected	mFA22	Selected
fMO21	Selected	mKG22	Selected
fSB21	Selected	mMK22	Selected
fFH22	Selected	mMM22	Selected
fLB22	Excluded: stutter	mMR22	Excluded: stutter
fOM22	Selected	mPP22	Selected
fTH22	Selected	mSJ22	Selected
fCW23	Selected	mSM22	Selected
fEW23	Selected	mTR22	Selected
fKL23	Selected	mGB23	Selected
fKW23	Selected	mHB23	Selected
fMK23	Selected	mML23	Selected
fNR23	Excluded: interruptions due to laughter	mMR23	Selected
fNS23	Selected	mPK23	Selected
fTR23	Selected	mSB23	Selected
fAN25	Excluded: stutter	mTQ23	Selected
fFS25	Selected	mAH24	Selected
fKK25	Excluded: stutter	mEF24	Selected
fKS25	Selected	mPR24	Selected
fMB25	Selected	mTT24	Selected
fST25	Selected	mAH25	Selected
fKH60	Selected	mSK25	Selected
fSW60	Selected	mBT60	Selected
fBP61	Selected	mDD61	Selected
fSH61	Excluded: dialect	mGG61	Excluded: dialect
fBA62	Selected	mWK62	Selected
fUJ62	Selected	mJM63	Excluded: dialect
fBD63	Selected	mKA63	Excluded: dialect
fCW63	Selected	mUT63	Excluded: dialect
fRG63	Selected	mDW65	Selected
fDG64	Excluded: dialect	mJL65	Excluded: dialect
fDH65	Selected	mTK65	Selected
fIM66	Selected	mHK66	Selected
fCT68	Selected	mJB66	Excluded: dialect
fBS69	Excluded: lisp and dialect	mJR66	Selected
fHR69	Selected	mHP67	Excluded: dialect
fIP69	Selected	mVS67	Selected
fKR69	Selected	mHD68	Excluded: dialect
fSR69	Selected	mLH68	Selected
fKF70	Selected	mRM68	Selected
fMJ70	Selected	mWG68	Selected

fSS70	Selected	mWH68	Selected
fUV70	Selected	mPK70	Selected
fGS71	Selected	mVL70	Excluded: dialect
fRP71	Excluded: dialect	mBB71	Excluded: dialect
fHM73	Selected	mJW71	Selected
fUZ73	Selected	mPH71	Selected
fEU74	Excluded: stutter	mVE71	Selected
fAB75	Selected	mUR73	Excluded: lisp
fHL77	Excluded: dialect	mMT75	Selected
		mMV77	Selected
		mRH81	Selected

Notes: The first letter of the speaker ID indicates the gender, and the last two numbers indicate the speaker's age. There were voice recordings of 59 female and 61 male speakers. For the vocal stimuli set of the study, voice recordings of 12 female and 12 male speakers were excluded.

Appendix E. Confirmatory Factor Analysis Aaker's (1997) Brand Personality Scale

Construct/ Item	Standardized Factor Loading (|≥.5|)	CR (>.7)	AVE (>.5)
Competence			
confident	.77***		
corporate	.66***		
hard working	.38***		
intelligent	.76***		
leader	.66***	.86	.43
reliable	.71***		
secure	.68***		
successful	.79***		
technical	.33***		
Excitement			
contemporary	.73***		
cool	.74***		
daring	.48***		
exciting	.68***		
imaginative	.66***		
independent	.64***	.90	.44
spirited	.65***		
trendy	.77***		
unique	.61***		
up-to-date	.78***		
young	.50***		
Ruggedness			
rugged	.27***		
masculine	.00	.15	.07
outdoorsy	.42***		
tough	-.08***		
western	.28***		
Sincerity			
cheerful	.59***		
down-to-earth	.63***		
family-oriented	.70***		
friendly	.72***		
honest	.71***		
original	.51***	.88	.41
real	.68***		
sentimental	.67***		
sincere	.75***		
small town	.16***		
wholesome	.70***		
Sophistication			
charming	.77***		
feminine	.12***		
glamorous	.55***	.74	.35
good looking	.73***		
smooth	.54***		
upper class	.62***		

Notes: Method: Maximum Likelihood Estimation with Robust Standard Errors and Satorra-Bentler (S-B) correction. Values below the specified thresholds are highlighted in red. Fornell-Lacker criterion not met (AVE > squared correlation of the latent construct with the discriminant construct). N = 3,945 individual voice ratings; *** $p < .001$.

Global Model Fit Indices: S-Bχ^2 (809) = 27,175.68 ($p < .001$); χ^2/df = 33.59; Standardized Root Mean Square Residual (SRMR) = .09; Root Mean Square Error of Approximation (RMSEA) = .09; Comparative Fit Index (CFI) = .71; Tucker-Lewis Index (TLI) = .69; Goodness-of-Fit Index (GFI) = .68; Adjusted Goodness of Fit Index (AGFI) = .65; Normed Fit Index (NFI) = .70.

Abbreviations: CR = Composite Reliability; AVE = Average Variance Extracted.

Appendix F. Confirmatory Factor Analysis Geuens et al.'s (2009) Brand Personality Scale

Construct/ Item	Standardized Factor Loading (|≥.5|)	CR (>.7)	AVE (>.5)
Activity			
dynamic	.76***	.99	.57
innovative	.72***		
active	.77***		
Simplicity			
simple	34.05***	475.17	579.74
ordinary	.01***		
Responsibility			
down-to-earth	.56***	.71	.46
responsible	.76***		
stable	.69***		
Emotionality			
romantic	.81***	.70	.53
sentimental	.64***		
Aggressiveness			
bold	.83***	.59	.44
aggressive	.43***		

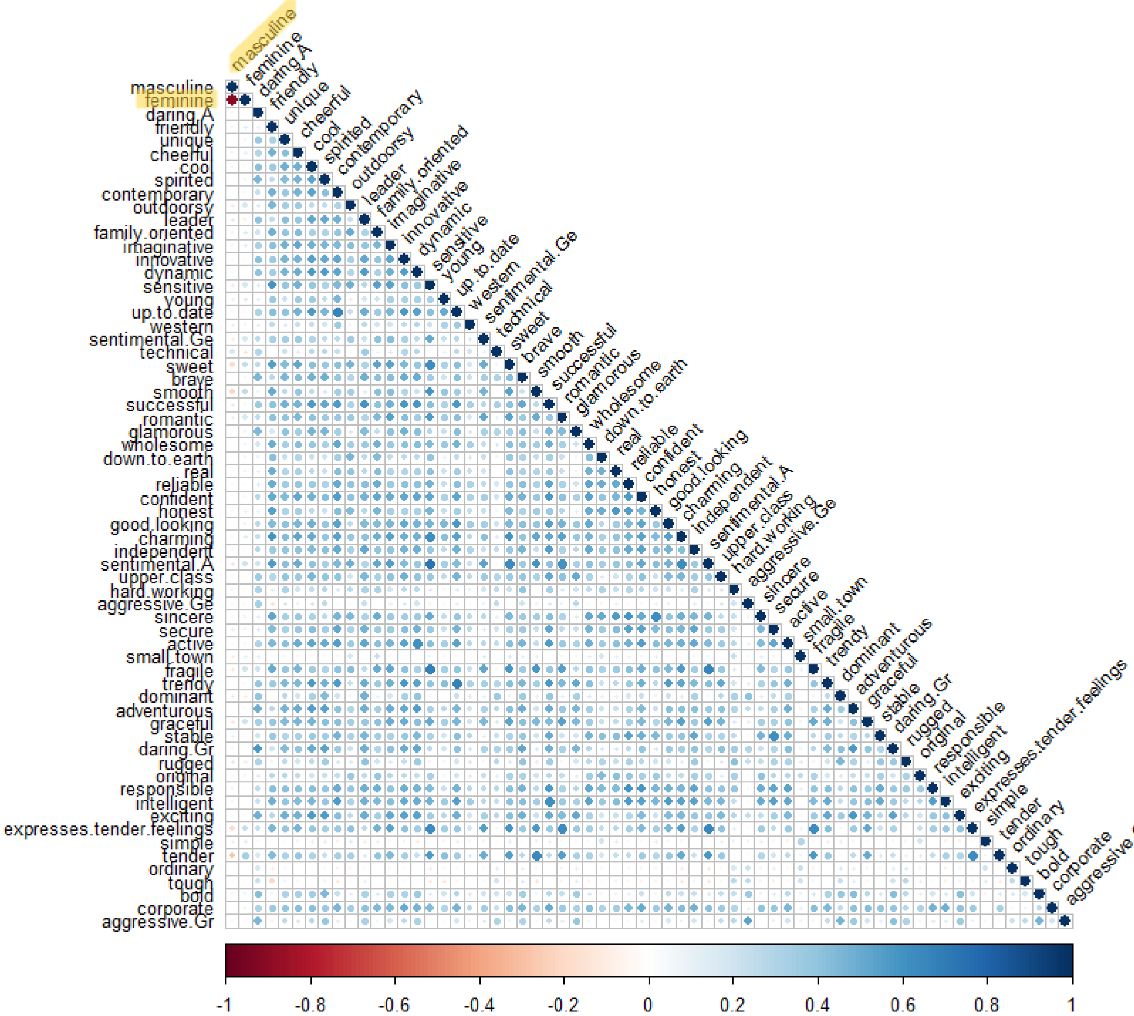
Notes: Method: Maximum Likelihood Estimation with Robust Standard Errors and Satorra-Bentler (S-B) correction. Values below the specified thresholds are highlighted in red. Fornell-Lacker criterion not met (AVE > squared correlation of the latent construct with the discriminant construct). N = 3,945 individual voice ratings; *** p < .001.
Global Model Fit Indices: S-Bχ^2 (44) = 1,020.03 (p < .001); χ^2/df = 23.18; Standardized Root Mean Square Residual (SRMR) = .05; Root Mean Square Error of Approximation (RMSEA) = .08; Comparative Fit Index (CFI) = .92; Tucker-Lewis Index (TLI) = .88; Goodness-of-Fit Index (GFI) = .96; Adjusted Goodness of Fit Index (AGFI) = .93; Normed Fit Index (NFI) = .92.
Abbreviations: CR = Composite Reliability; AVE = Average Variance Extracted.

Appendix G. Confirmatory Factor Analysis Grohmann's (2009) Brand Personality Scale

Construct/ Item	Standardized Factor Loading (|≥.5|)	CR (>.7)	AVE (>.5)
Feminity			
expresses tender feelings	.79***	.89	.58
fragile	.79***		
graceful	.67***		
sensitive	.80***		
sweet	.76***		
tender	.73***		
Masculinity			
adventurous	.53***	.80	.41
aggressive	.70***		
brave	.77***		
daring	.56***		
dominant	.50***		
rugged	.79***		

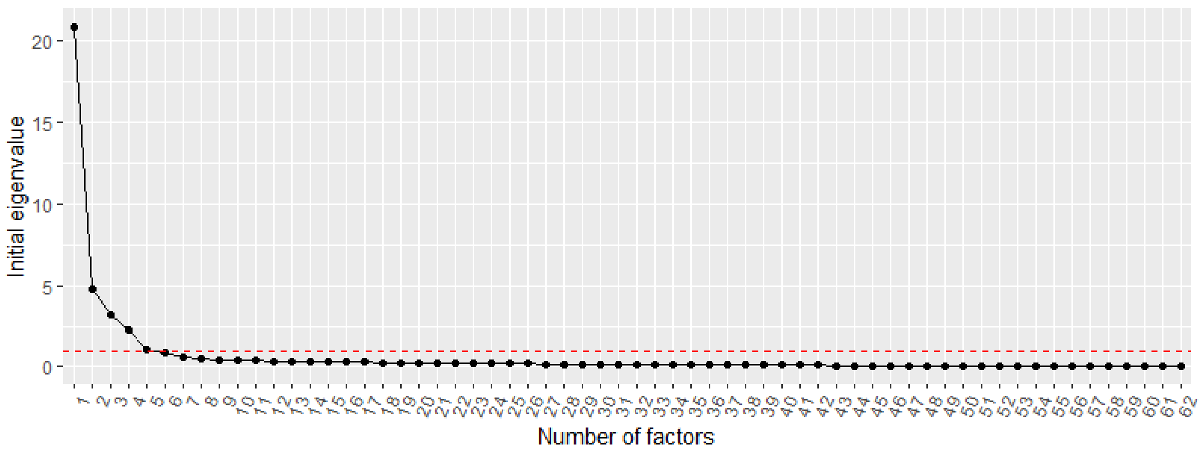
Notes: Method: Maximum Likelihood Estimation with Robust Standard Errors and Satorra-Bentler (S-B) correction. Values below the specified thresholds are highlighted in red. Fornell-Lacker criterion not met (AVE > squared correlation of the latent construct with the discriminant construct). N = 3,945 individual voice ratings; *** p < .001.
Global Model Fit Indices: S-Bχ^2 (53) = 2,178.384 (p < .001); χ^2/df = 41.1; Standardized Root Mean Square Residual (SRMR) = .09; Root Mean Square Error of Approximation (RMSEA) = .10; Comparative Fit Index (CFI) = .90; Tucker-Lewis Index (TLI) = .87; Goodness-of-Fit Index (GFI) = .91; Adjusted Goodness of Fit Index (AGFI) = .87; Normed Fit Index (NFI) = .90.
Abbreviations: CR = Composite Reliability; AVE = Average Variance Extracted.

Appendix H. Correlogram of Personality Traits



Notes: Positive correlations are displayed in blue, and negative correlations in red. The color intensity and the size of the circle are proportional to the correlation coefficients. Unsignificant correlation coefficients (p > .05) are blank.
Items ending with ".A", ".Ge", and ".Gr" indicate from which brand personality scale the item is from (A = Aaker, Ge = Geuens et al., Gr = Grohman). For example, "aggressive.Gr" means that this "aggressive" refers to the Grohman (2009) brand personality scale item, while "aggressive.Ge" refers to the Geuens et al. (2009) scale item. In German, these were two distinct items.

Appendix I. Scree Plot of Unrotated Factors



Notes: The red dotted line indicates the "elbow" of the plot, which represents the threshold for retaining the initial factors extracted from the observed variables that maximize the variance accounted for (Eigenvalues > 1.0).

Appendix J. Exploratory Factor Analysis (1st rotation)

Factor/ Item	Standardized Factor Loading (|≥.5|)	Indicator Reliability (≥.4)	Communality (≥.4)	Item-Total Correlation (≥.4)	Cronbach's Alpha (>.7)	Eigenvalue (>1.0)
Factor 1					.89	4.18
wholesome	.54	.29	.51	.68		
down-to-earth	.62	.38	.42	.61		
real	.64	.41	.46	.66		
reliable	.70	.49	.59	.77		
honest	.66	.43	.54	.72		
sincere	.69	.47	.58	.75		
secure	.61	.37	.52	.69		
stable	.66	.43	.50	.68		
responsible	.57	.33	.52	.69		
Factor 2					.93	5.64
imaginative	.51	.26	.50	.66		
sensitive	.71	.50	.67	.79		
sentimental (Geuens et al., 2009)	.73	.54	.50	.68		
sweet	.66	.44	.58	.75		
smooth	.68	.46	.57	.71		
romantic	.70	.49	.59	.76		
sentimental (Aaker, 1997)	.75	.56	.69	.83		
fragile	.73	.53	.63	.79		
graceful	.50	.25	.50	.66		
expresses tender feelings	.80	.64	.66	.81		
tender	.74	.54	.61	.74		
Factor 3					.88	4.25
daring (Aaker, 1997)	.68	.46	.49	.68		
spirited	.59	.35	.51	.70		
leader	.59	.35	.57	.71		
glamorous	.55	.30	.50	.63		
dominant	.69	.48	.49	.64		
adventurous	.58	.34	.54	.71		
daring (Grohman, 2009)	.68	.46	.56	.75		
rugged	.52	.27	.47	.49		
exciting	.57	.32	.57	.67		
aggressive (Grohman, 2009)	.69	.47	.42	.54		
Factor 4					.65	1.17
young	.65	.42	.49	.61		
up-to-date	.56	.32	.61	.61		

Notes: Method: Principal Axis Factoring (PAF). **Rotation**: Oblimin with Kaiser Normalization. n = 1,970 individual voice ratings. **Global Model Fit Indices**: Root Mean Square Error of Approximation (RMSEA) = .04; Standardized Root Mean Square Residual (SRMR) = .02; Tucker-Lewis Index (TLI) = .95.

Appendix K. Exploratory Factor Analysis (Final Result)

Factor/ Item	Standardized Factor Loading (|≥.5|)	Indicator Reliability (≥.4)	Communality (≥.4)	Item-Total Correlation (≥.4)	Cronbach's Alpha (>.7)	Eigenvalue (>1.0)
Factor 1					.89	3.99
wholesome	.58	.34	.51	.68		
real	.65	.42	.46	.66		
reliable	.74	.54	.59	.77		
honest	.68	.46	.54	.72		
sincere	.70	.49	.58	.75		
secure	.68	.46	.52	.69		
stable	.70	.49	.50	.68		
responsible	.62	.38	.52	.69		
Factor 2					.93	5.82
imaginative	.52	.27	.50	.66		
sensitive	.69	.48	.67	.79		
sentimental (Geuens et al., 2009)	.67	.44	.50	.68		
sweet	.65	.42	.58	.75		
smooth	.76	.58	.57	.71		
romantic	.73	.53	.59	.76		
sentimental (Aaker, 1997)	.74	.54	.69	.83		
fragile	.72	.51	.63	.79		
graceful	.54	.29	.50	.66		
expresses tender feelings	.79	.63	.66	.81		
tender	.83	.68	.61	.74		
Factor 3					.88	4.07
daring (Aaker, 1997)	.69	.47	.49	.68		
spirited	.61	.37	.51	.70		
leader	.60	.36	.57	.71		
glamorous	.58	.34	.50	.63		
dominant	.66	.44	.49	.64		
adventurous	.61	.37	.54	.71		
daring (Grohman, 2009)	.69	.47	.56	.75		
exciting	.60	.36	.57	.67		
aggressive (Grohman, 2009)	.69	.47	.42	.54		

Notes: Method: Principal Axis Factoring (PAF). **Rotation**: Oblimin with Kaiser Normalization. n = 1,970 individual voice ratings. **Global Model Fit Indices**: Root Mean Square Error of Approximation (RMSEA) = .04; Standardized Root Mean Square Residual (SRMR) = .02; Tucker-Lewis Index (TLI) = .95.

Appendix L. Assessment of Univariate & Multivariate Normality

Factor/ Item	Min	Max	Skewness (|<2|)	p-value (>.05)	Kurtosis (|<7|)	p-value (>.05)
Factor 1						
wholesome	1.0	5.0	-.43	.001	-.30	.001
real	1.0	5.0	-.71	.001	-.01	.001
reliable	1.0	5.0	-.37	.001	-.40	.001
honest	1.0	5.0	-.64	.001	.01	.001
sincere	1.0	5.0	-.52	.001	-.24	.001
secure	1.0	5.0	-.16	.001	-.89	.001
stable	1.0	5.0	-.23	.001	-.67	.001
responsible	1.0	5.0	-.11	.001	-.73	.001
Factor 2						
imaginative	1.0	5.0	.32	.001	-.84	.001
sensitive	1.0	5.0	-.11	.001	-.88	.001
sentimental (Geuens et al., 2009)	1.0	5.0	.42	.001	-.74	.001
sweet	1.0	5.0	.05	.001	-.98	.001
smooth	1.0	5.0	-.05	.001	-.96	.001
romantic	1.0	5.0	.66	.001	-.52	.001
sentimental (Aaker, 1997)	1.0	5.0	-.06	.001	-.98	.001
fragile	1.0	5.0	-.07	.001	-.86	.001
graceful	1.0	5.0	.47	.001	-.59	.001
expresses tender feelings	1.0	5.0	.39	.001	-.92	.001
tender	1.0	5.0	.38	.001	-.96	.001
Factor 3						
daring (Aaker, 1997)	1.0	5.0	.94	.001	.35	.001
spirited	1.0	5.0	.80	.001	-.06	.001
leader	1.0	5.0	.43	.001	-.62	.001
glamorous	1.0	5.0	1.48	.001	1.89	.001
dominant	1.0	5.0	.90	.001	-.02	.001
adventurous	1.0	5.0	.64	.001	-.39	.001
daring (Grohman, 2009)	1.0	5.0	.64	.001	-.24	.001
exciting	1.0	5.0	.89	.001	.09	.001
aggressive (Grohman, 2009)	1.0	5.0	1.62	.001	2.13	.001
Multivariate			14,280.52	.001	109.15	.001

Notes: p-values < .05 indicate that univariate and multivariate normal distribution did not hold for the present data, although skewness and kurtosis values are within normal distribution thresholds of |<2| and |<7|, respectively (West et al., 1995). Multivariate normality was assessed using Mardia's coefficient (West et al., 1995).

Appendix M. Pairwise Correlation Coefficients between the Acoustical Measures

	Speaking Rate	Average Silent Pause Duration	Silent Pause Frequency per Minute	Articulation Rate	Intensity Variability	Mean f0	SD f0	Range f0	h1-h2	Spectral Slope	Spectral Tilt	HNR	CPPS	Jitter	Shimmer
Speaking Rate	1.00														
Average Silent Pause Duration	.28	1.00													
Silent Pause Frequency per Minute	-.45	.60	1.00												
Articulation Rate	.29	-.03	.10	1.00											
Intensity Variability	-.44	.19	.54	-.15	1.00										
f0 Mean	-.11	.16	.13	-.09	.14	1.00									
f0 SD	-.09	.05	.06	-.03	.05	**.82**	1.00								
f0 Range	.07	-.22	-.18	.10	-.23	.24	.65	1.00							
h1-h2	-.09	.14	.18	.11	.06	.38	.22	-.15	1.00						
Spectral Slope	.14	-.00	-.20	-.25	.04	.15	.06	-.10	-.08	1.00					
Spectral Tilt	-.15	-.21	-.02	.19	-.15	.14	.19	.18	.01	-.40	1.00				
HNR	-.30	.14	.18	-.29	.38	**.85**	.63	.06	.37	.02	.06	1.00			
CPPS	-.19	.00	-.02	-.46	.47	.13	-.02	-.27	-.10	.53	-.35	.40	1.00		
Jitter	.32	-.13	-.19	.37	-.49	-.73	-.47	.09	-.24	-.21	.01	**-.90**	-.59	1.00	
Shimmer	.31	-.07	-.06	.44	-.33	-.65	-.43	.08	-.27	-.19	.04	**-.88**	-.57	**.87**	1.00

Notes: Correlation coefficients greater than |.8| indicate a strong pairwise correlation and are highlighted in red.
Abbreviations: CPPS = smoothed cepstral peak prominence; f0 = fundamental frequency; HNR = harmonics-to-noise ratio; SD = standard deviation.

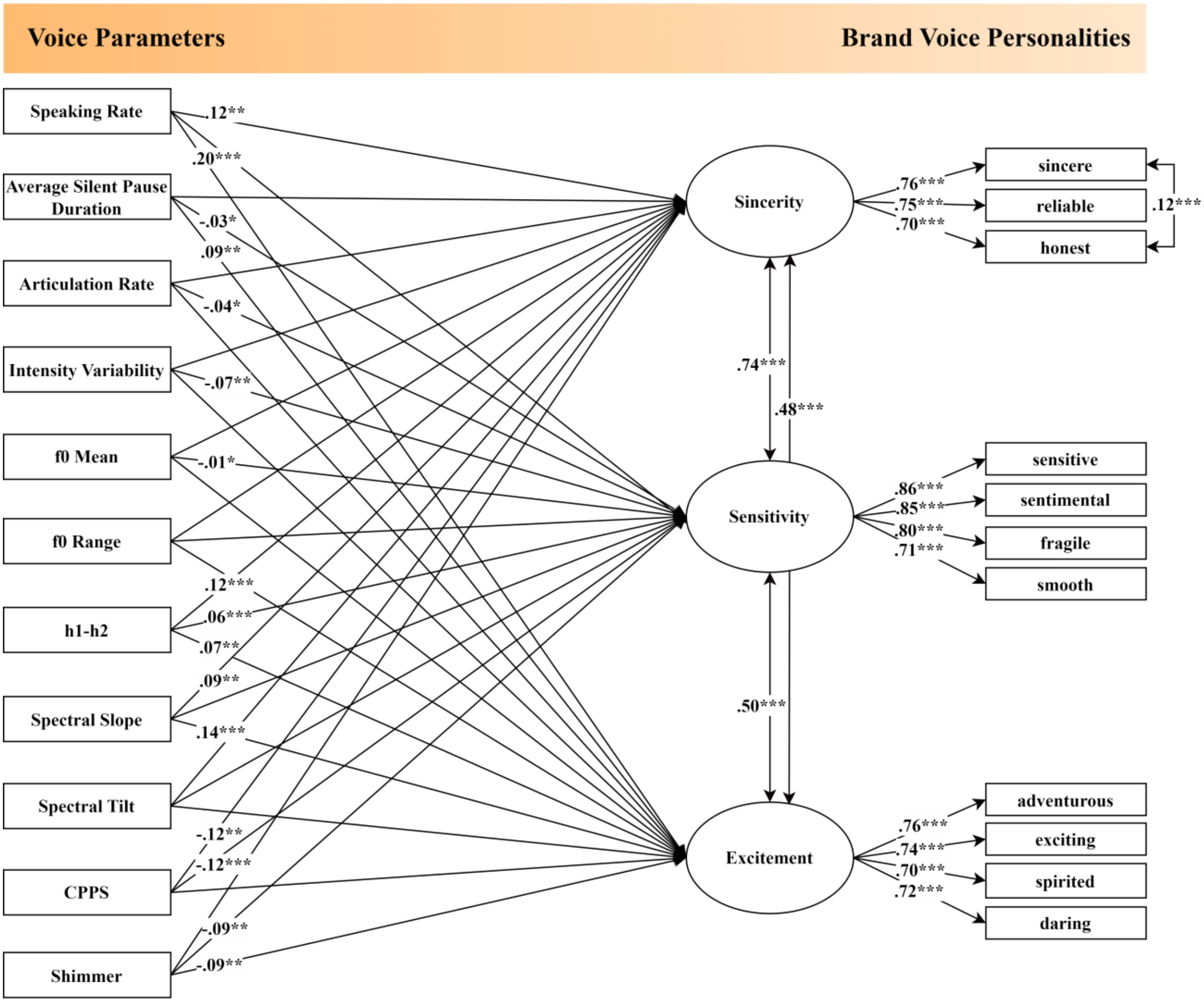
Appendix N. Initial Multicollinearity Assessment with 14 Acoustical Measures

Acoustical Measures	R^2	Collinearity Statistics	
		Tolerance > .2	**VIF < 5**
Speaking Rate	.88	.12	**8.35**
Average Silent Pause Duration	.89	.12	**8.96**
Silent Pause Frequency per Minute	.91	.09	**11.71**
Articulation Rate	.61	.39	2.58
Intensity Variability	.61	.39	2.55
f0 Mean	.91	.09	**10.91**
f0 SD	.91	.09	**10.93**
f0 Range	.78	.22	4.46
h1-h2	.31	.69	1.46
Spectral Slope	.51	.49	2.05
Spectral Tilt	.34	.66	1.51
CPPS	.77	.23	4.34
Jitter	.88	.12	**8.35**
Shimmer	.82	.18	**5.46**

Notes: VIF greater than 5 indicates high (multi)collinearity and are highlighted in red. VIF = $1/(1- R^2)$; Tolerance = $1 – R^2$.
Abbreviations: CPPS = smoothed cepstral peak prominence; f0 = fundamental frequency; SD = standard deviation; VIF = variance inflation factor.

Appendix O. Brand Voice Personality Models (BVP-Models)

A. BVP-Model for Female Voices

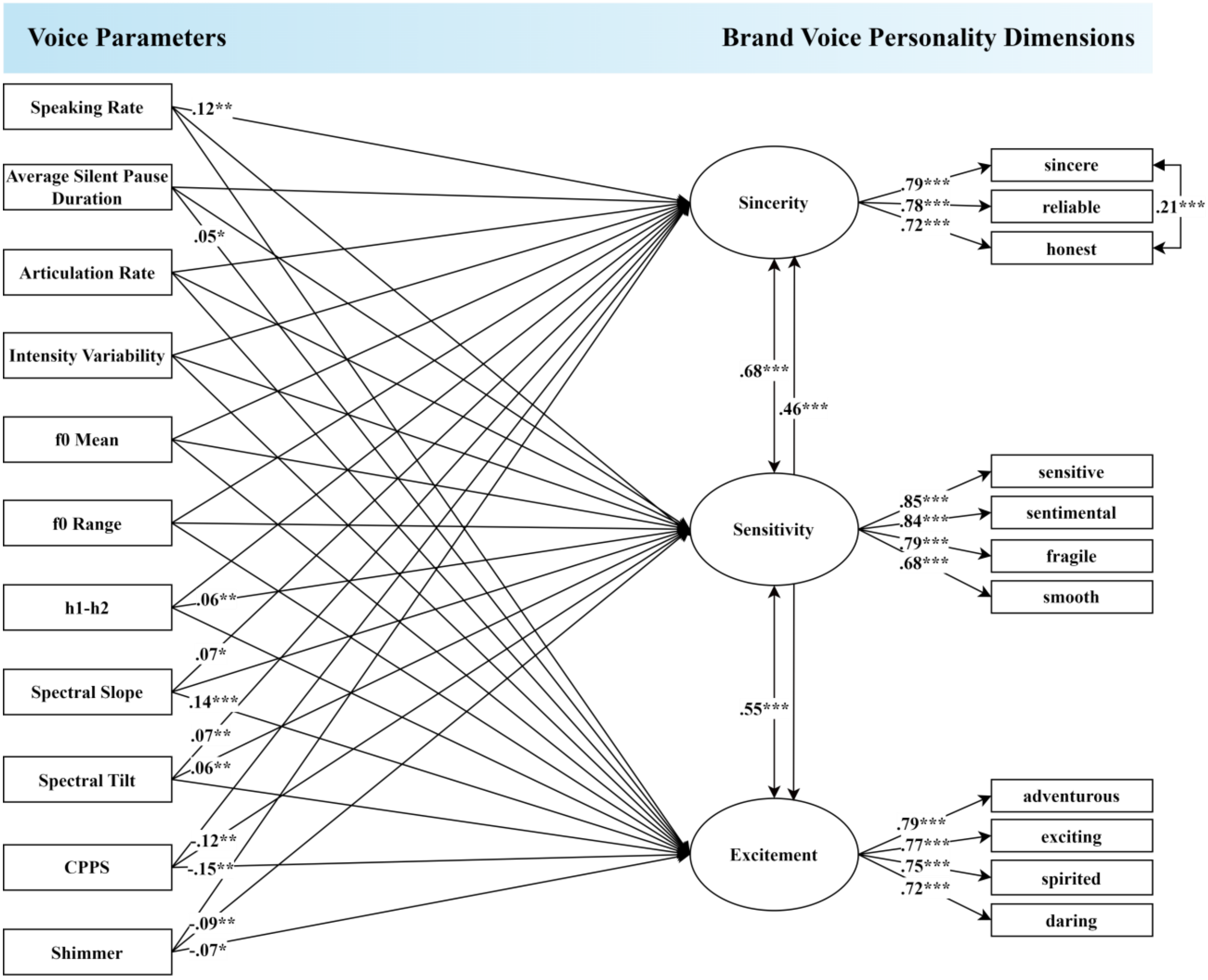


Notes: This figure is simplified to provide a better overview of the model: (1) only the significant paths are displayed; (2) socio-demographic and survey-related data are included as covariates in the analysis but are not shown in this model but displayed below. *p < .1, **p < .05, ***p < .001.

Significant Covariates:
Listener's Gender → Sensitivity: β = .10***
Listener's Age → Sensitivity: β = .06**
Listener's Income → Sensitivity: β = -.06**

Listener's Gender → Excitement: β = .11***
Listener's Age → Excitement: β = .06**
Listener's Education → Excitement: β = .06**

B. BVP-Model for Male Voices



Notes: This figure is simplified to provide a better overview of the model: (1) only the significant paths are displayed; (2) socio-demographic and survey-related data are included as covariates in the analysis but are not shown in this model but displayed below. *p < .1, **p < .05, ***p < .001.

Significant Covariates:
Listener's Gender → Sincerity: β = -.08**

Listener's Gender → Sensitivity: β = .05*
Listener's Income → Sensitivity: β = -.05*

Listener's Gender → Excitement: β = .08**
Listener's Age → Excitement: β = -.22***
Listener's Education → Excitement: β = -.07**

Essay 3

Brand Voice Across Borders: A Comparison of German and American Brand Personality Perceptions Through Voice

Author: Olga Bosak

Abstract

With regard to voice-based technology, international acting companies must demonstrate an understanding of the significant impact of culture as an acoustic cue and consider adapting their brand's voice to align with the cultural context in which they intend to interact with customers. The perception of personality traits in voices differs according to cultural context, with listeners from diverse cultural groups attributing distinctive personality characteristics to the same speaker. Despite the efforts to develop culturally sensitive voice assistants, there is a limited body of research and adaptation in brand personality perception across languages. This study addresses this gap by investigating the vocal perception of brand personalities in the United States and comparing the findings with those of German research. The cross-cultural comparison demonstrates that listeners in both cultures perceive the exact brand personality dimensions of sincerity, sensitivity, and excitement through voice alone. In contrast, a comparison of the developed American and German brand voice personality models (BVP-Models), that determine which voice parameters induce the perception of brand voice personality dimensions, reveals that the brand voice profiles exhibit more differences than similarities. The findings support the theory that cross-cultural variations in learned social interaction norms and rules are a significant determinant of vocal brand personality perception. The study emphasizes the importance of organizations with an international presence in selecting or designing a brand voice that is adapted to the cultural norms of their customers.

Keywords: brand personality, brand voice, personality perception, cross-cultural research, voice AI, voice marketing, replication

1. Introduction

As voice-based interactions in the form of voice assistants gain prominence and become increasingly prevalent, it becomes crucial for companies to include voice brand touchpoints and actively engage in this area (Seaborn et al., 2022; 2024). Voice artificial intelligence (AI) – AI technologies that can perform speech recognition and natural language processing (NLP) to engage in natural dialogues with users such as voice assistants (Wang et al., 2023) – enables individuals to control home settings, engage with companies and brands for assistance and information, or place orders (Vixen Labs & Digitalscouting, 2022). Current examples of utilizing voice in its marketing strategy are, for instance, Nike's voice-activated shoe assistant on Siri and Google Home (Schwartz, 2019), Volkswagen's voice assistant "IDA" (Volkswagen Group, 2024), or Coca-Cola's voice-activated campaigns via Alexa rewarding users with coupons or free samples (Schwartz, 2020; VML, 2023). Further, the emergence of sophisticated technological innovations and ubiquitous digitalization have led to improved global connectivity, facilitating and accelerating access to new markets. This worldwide connectivity is forcing companies to undergo digital transformations to survive in competitive markets or expand their operations on an international scale (Hess et al., 2020).

While an expansion into international markets offers opportunities, it also presents several risks and challenges. One such challenge is the communication of a consistent brand identity, which requires companies to understand their customers' cultural and social backgrounds and adjust their communication strategies for various audiences (Chernatony et al., 1995). When it comes to voice AI, such as conversational agents and voice assistants, international acting companies must comprehend the noteworthy influence of culture as an acoustic cue and consider adapting their brand's voice to align with the cultural context in which they wish to interact with customers. The way a brand is represented through its voice influences the perception of the brand personality and, therefore, the brand identification (Nam et al., 2011). This perception depends not only on the voice and its acoustic characteristics but also on the cultural context of the customer (Krauss et al., 2002).

Cross-cultural perceptual studies have revealed the influence of sociocultural factors on the perception of personalities through voice (H. O. Lee & Boster, 1992; Peng et al., 1993; van Bezooijen, 1995). Vocal stereotypes can vary between cultures due to differences in how personality traits are learned and perceived from speech (Kreiman & Sidtis, 2011; Waaramaa et al., 2021). For instance, Japanese cultural norms dictate that women speak more politely than men. A higher pitch in the voice conveys this politeness. Although women naturally have medium-pitched voices, cultural norms have conditioned Japanese women to speak in a higher pitch (Krauss et al., 2002). A study by van Bezooijen (1995) demonstrated the influence of this cultural conditioning on voice perceptions and preferences by showing that in Japan, high-pitched voices in women are perceived as more attractive, whereas in the Netherlands, a medium or low-pitched voice is preferred. In a communication system that relies solely on voice, it is therefore crucial for organizations to understand the nuances of a culture, particularly when the objective is to enhance the user experience (UX) and build meaningful customer relationships internationally (Pang, 2021).

Despite efforts to provide culturally sensitive voice assistants in various languages, there is a limited body of research and adaptation in (brand) personality perception across languages. Although Alexa can imitate human emotions through excited or disappointed voice tones (available in British English and Japanese), it is not yet possible for any voice assistant to reflect a specific personality by adjusting its voice (Wenden, 2020). Building on previous research examining the vocal perception of personalities in different language contexts, this study investigates the role of cultural brand personality perception through the voice. To that end, a perceptual study is conducted in the United States, and the results are compared with those of the previous research on the vocal perception of brand personalities conducted in Germany.

The number of users of voice assistants in the United States has increased steadily over the past two years, from 142 million in 2022 to 149.8 million in 2024, with projections indicating a further increase to 157.1 million users by 2026 (eMarketer). The United States exhibits the highest prevalence of individuals owning any smart speaker (35% in 2022), surpassing Canada (30%) and Australia (28%; Triton Digital, 2022). Furthermore, 57% of individuals in the United States reported utilizing voice assistants at least twice or three times per month, a figure that is comparable to the 56% of individuals in the United Kingdom and higher than the 50% of individuals in Germany who reported the same frequency of use in 2022 (Vixen Labs & Digitalscouting, 2022). Given that the United States represents one of the largest markets for smart speakers, it is imperative to research the vocal perception of brand personality in American English. However, such research currently needs to be completed.

This study conducts a perceptual study in American English, following the recommendations of the two preceding essays on cross-cultural research in vocal brand personality perception. It replicates and extends the analysis presented in Essay 2, which empirically examined how brand personalities are perceived through voice in a German-speaking context. Following the German approach, the present study's *first objective* is determining which brand personalities can be perceived through voice in an American English-speaking context by developing an American brand voice personality scale (BVP-Scale). Further, the study's *second objective* is developing an American brand voice personality model (BVP-Model) that determines how female and male voices induce the perception of brand voice personality dimensions. The *third objective* is determining cultural differences in vocal brand personality perceptions between the United States and Germany. Therefore, the German and American BVP-Scales and BVP-Models are compared.

The value of this cross-cultural study lies in its capacity to assess the generalizability of prior empirical findings and to contribute to a better understanding of the cultural influence on vocal brand personality perception. The findings will assist companies operating in the U.S. and internationally in selecting and designing their brand voices to enhance the user experience in voice-based interactions and build customer trust, loyalty, and engagement with the brand (Ha & Stoel, 2009).

In addition to the managerial recommendations, the findings contribute to the field of vocal personality perception and voice marketing research in three ways. First, they contribute to the limited research on the vocal perception of brand personalities, as previous research focused on examining human personalities. Second, this study is the first to investigate the vocal brand personality perception within a cross-cultural context. Third, this study's character of replication contributes to the applicability of the utilized methods and models to facilitate the development of global strategies on vocal brand personality perception.

The remaining sections are organized as follows: First, the theoretical background on speech, communication, personality perception, and culture will be provided to clarify key concepts and terms used in this study and to situate the work with existing knowledge and theories. Subsequently, the present study's research aim, scope, and framework will be explained in greater detail. Next, a description of the online survey methodology, which was employed to determine which brand personality traits are associated with voices, will be provided. After that, the results on vocal brand personality perception in American English will be presented, and the American English findings will be considered and compared cross-culturally with those from Germany in the discussion. The study will conclude with an overview of the managerial and theoretical implications and recommendations for future research.

2. Background

Due to this study's cross-cultural nature, the following existing knowledge and theories are presented to help understand the national specifics of vocal personality perception and trans-national communication patterns. In doing so, the similarities and differences between Germany and the United States are pointed out, as the study focuses on comparing these two cultures concerning the vocal perception of brand personality.

2.1. Speech, Communication, and Culture

Human speech provides information from at least two sources: a verbal channel, which encodes the semantic content of a message, and a paraverbal, i.e., a vocal channel, which conveys paralinguistic information through variations in voice parameters, such as speaking rate, intonation, or pitch (Apple et al., 1979). Individual variations in voice parameters around a mean determine voice uniqueness, which gives each person an individual vocal signature (Belin et al., 2011; Rodero, 2013). Listeners interpret this paralinguistic information and derive judgments about the speaker that go beyond the content of a message, such as the speaker's gender, age, regional origin, social class, identity, and personality (Klasmeyer & Sendlmeier, 1997; Zäske et al., 2020).

The anatomy of the vocal tract limits the variety of sounds the voice can produce. However, cultural norms influence how people use their voices, with certain cultures establishing rules on speaking styles that are perceived as desirable for particular groups of speakers (Krauss et al., 2002; Waaramaa et al., 2021). This observation can be illustrated using a speaker's pitch range, which is anatomically constrained by the length of the vocal tract. However, cultural norms influence the specific location within this range at which female and male speakers place their voices. The description above of women in Japan conditioned to speak with a high pitch to be perceived as polite and preferred in general serves to demonstrate this phenomenon (van Bezooijen, 1995).

Culture is defined as “the collective programming of the mind that distinguishes the members of one group […] of people from others” (Hofstede, 2011, p. 3). In addition to sharing values and norms, individuals within a culture also possess a common language, communication patterns, and speaking styles (Hogan & Bond, 2009). Verbal and nonverbal communication allows for expressing and perceiving emotions, attitudes, and personalities, which may differ between cultures depending on the social standards applied during voice production and perception (Krauss et al., 2002). Theoretical perspectives exist regarding the underlying concepts that determine national communication styles and explain possible transnational differences. The following discusses two central theories on cultural categories of communication. These include Hall's (1959, 1976) *concept of high/low context cultures* and Lewis's (2006) *model on cultural categories of communication*.

According to Hall (1959, 1976), cultures can be divided into high-context and low-context cultures on a continuum (see **Figure 1A**). In this case, context describes “the information that surrounds an event” (Hall & Hall, 2006, p. 6). In *high-context cultures*, the information in communications is not explicitly transmitted through the spoken word but through contextual and nonverbal cues, such as paraverbal features (Hall & Hall, 2006; Nishimura et al., 2008). In such cultures, communication is indirect, ambiguous, harmonious, and quiet and is strongly influenced by interpersonal relationships and social behavioral norms (Korac-Kakabadse et al., 2001; Nishimura et al., 2008). Furthermore, cultures characterized as high-context are predominantly collectivist cultures where group identification, harmony, and a “we” consciousness are important (Hofstede, 2011; Korac-Kakabadse et al., 2001). Most high-context cultures are located in Asia, including countries such as Japan, China, and Indonesia (Hall & Hall, 2006).

Additionally, several European countries, including Greece and Spain, are classified as high-context cultures (Hall & Hall, 2006).

In contrast, individuals in *low-context cultures* tend to communicate information explicitly through messages, and the context is often neglected (Hall, 1976). Low-context communication is direct, clear, open, and lively, focusing on efficiency instead of a harmonious interaction (Korac-Kakabadse et al., 2001). Low-context cultures are predominately individualistic, in which individuals focus on independence, privacy, personal preferences, and freedom (Hofstede, 2011; Korac-Kakabadse et al., 2001). Most Western countries, including the United States, Canada, New Zealand, and European countries, especially the German-speaking countries, are low-context cultures (Hall & Hall, 2006).

Similar to Hall's (1959, 1976) two-dimensional categorization of cultures, Lewis (2010) organized cultures into three communication categories: linear-active, reactive, and multi-active (see **Figure 1B**). The communication style of *linear-active cultures* can be described as direct and polite, with an equal division of speaking and listening and a focus on facts over emotions and body language. This category is comparable to the low-context culture and comprises English-speaking countries, like the United States, Australia, the U.K., Scandinavia, and German-speaking countries. Individuals of *reactive cultures* are good listeners and polite, indirect, and patient speakers. They value diplomacy, subtle body language, and cultivating good relationships. This category is comparable to the high-context culture and comprises Asian countries such as Japan, Vietnam, China, and South Korea. *Multi-actives* are impulsive, emotional, warm, impatient, and communicative. They engage in simultaneous speech and listening, which frequently results in the interruption of ongoing conversations. Additionally, they exhibit expressive body language. This category encompasses countries in the African, Arabic, Slavic, and Spanish-speaking regions of Europe, as well as North and South America. (Korac-Kakabadse et al., 2001; Lewis, 2006; Nishimura et al., 2008)

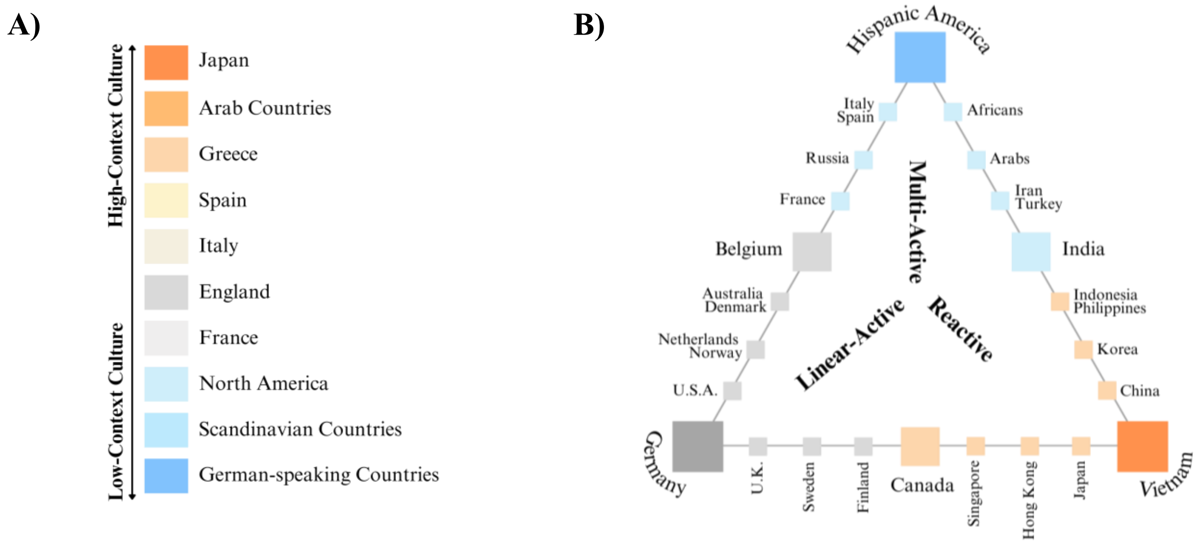


Figure 1. Cultures Categorized According to: A) High/Low Context Cultures; B) Linear-active, Reactive, Multi-active Cultures
Notes: Own illustration based on A) Hall & Hall (2006); B) Lewis (2006).

Concerning the two communication models described, the communication patterns of German and American English are similar. Both cultures are categorized as low-context cultures with a linear-active communication style. Additionally, both countries are individualistic, and both languages belong to the Indo-European language family, specifically the branch of West Germanic languages, which contributes to the observed similarities (Paul Roberge, 2020). Communication in the two cultures is characterized by directness, politeness, factual focus, and a lack of emotion. In a direct comparison, Germans are more direct, explicit, and self-referenced and behave more impolitely when making requests than Americans (House, 2018; House & Kasper, 1981).

Comparing paraverbal voice parameters between German and American English reveals that major differences do not exist. Coupé et al. (2019) examined 17 languages and nine language families on speech and information rate, with the result that the German average speech rate of 4-8 syllables per second is slightly slower than the English speech rate of 5-8 syllables per second. This result is consistent with other studies on cross-linguistic comparisons in spontaneous speech. Fenk-Oczlon and Fenk (2010) and Pellegrino et al. (2011) found that German's average speech rate is 5.5 or 5.97 syllables per second, respectively, while English's average is 5.77 or 6.19. Concerning the intonation pattern of both languages, Americans show more variability and dynamism in intonation using a wider pitch range than Germans (Hirst & Di Cristo, 1998; Mennen et al., 2012). Further, a perceptual analysis of American and German speakers conducted by Scherer (1974) demonstrated that qualities in voice, such as breathiness or creakiness, were perceived similarly by expert judges and naïve listeners.

2.2. Vocal Personality Perception and Culture

The perception of personality traits in voices differs according to cultural context, with listeners from different cultural groups attributing distinctive personality characteristics to the same speaker (Kreiman & Sidtis, 2011). This observation can be attributed to the fact that the cultural patterns and social interaction rules that are learned and followed in one culture are not necessarily the same as those learned and followed in another culture (Brown et al., 1975; Hogan & Bond, 2009). For instance, D'Errico et al. (2013) explored the perception of charisma through the voices of Italian and French politicians. Among other findings, Italians were likelier to trust a speaker who used short pauses and a standard or higher pitch to convey charisma. At the same time, French participants preferred long pauses with a standard or lower pitch. Further, a fast speaking rate in male voices was negatively associated with trustworthiness in Korean but positively associated in American English (Charoenruk & Olson, 2018; H. O. Lee & Boster, 1992). The differences between collectivist and individualist cultures and the accompanying communication style differences in this latter example can be a significant influencing factor. In South Korea, a collectivist society, there is a greater emphasis on treating individuals with respect and communicating harmoniously than in individualistic cultures such as the United States. Consequently, trustworthy South Koreans consider their choice of words more carefully to avoid offending their counterparts, which results in a slower pace of speech.

Studies on differences in personality perceptions through voice between Germans and Americans are limited. According to Scherer (1979), extroverts are perceived as talking louder and with a more nasal voice. Additionally, American extroverts tend to make fewer pauses, while German extroverts produce more pauses than introverts (Scherer (1979) as cited in Mairesse and Walker, 2006). A low pitch influences the perception of competence in female voices in German (Krahé et al., 2021). In contrast, the results for American English indicated that female voices are perceived as more competent when the pitch is higher (Oksenberg et al., 1986). Furthermore, a slow articulation rate in female Germans is linked to increased levels of extraversion and neuroticism (Michalsky et al., 2020), whereas a fast speaking rate is found to improve perceptions of extraversion and neuroticism in female Americans (Addington, 1968; Aronovitch, 1976). Nevertheless, a direct comparison of these results is not entirely reliable, as the perceptual studies differ significantly in terms of stimulus (e.g., number of voices or content of speech), assessment methodology (external judgments or self-ratings), and study design (correlation study, cue synthesis or manipulation study).

In conclusion, differences in personality perception are primarily attributed to cross-cultural variations in learned communication patterns, social interaction norms, and rules. Therefore, it can be anticipated that, despite the comparable communication styles of German and American English, differences in vocal personality perception exist, which must be considered in intercultural voice marketing, particularly in brand voice selection. The existing perceptual studies on German and American English indicate that the timing dimension of speech, which incorporates voice parameters such as speaking rate, articulation rate, and pausing behavior, might be a decisive factor in cross-cultural differences in personality perception. This conclusion is based on the observation that most of the inconsistent findings in vocal personality perception (particularly those related to extraversion and neuroticism) were associated with the timing dimension of voice.

3. Research Framework

Extensive research exists on personality perception through voice, focusing on examining the impact of individual voice characteristics on personality trait judgments[1]. Nevertheless, two limitations exist in applying these findings to marketing and branding, especially in cross-cultural contexts:

1) Known perceptual studies investigated personality traits applicable to humans, not brands (Aronovitch, 1976; McAleer et al., 2014; Oksenberg et al., 1986). However, the characteristics of human personalities cannot be directly applied to brands in the same way. For example, the meaning of specific traits may differ significantly when transferred from the domain of human personalities to that of brands (Caprara et al., 2001). In addition, specific human personality traits do not necessarily apply to brands, and it is more probable that brands will be perceived with desirable traits that humans do not possess (Aaker, 1997; Azoulay & Kapferer, 2003). The perceptual study of Essay 2 was the first to employ brand personality traits in vocal personality perception, thereby addressing a significant research gap.

2) Limited studies have focused on cross-cultural differences in vocal personality perception (Chen et al., 2001; H. O. Lee & Boster, 1992; Peng et al., 1993; van Bezooijen, 1995). Moreover, most perceptual studies were conducted in American English-speaking contexts, as identified through the extensive literature review on personality perception through the voice of Essay 1. Of the 52 reviewed perceptual studies, 70% were conducted in American English. The remaining 30% were distributed across a total of 12 languages, which results in an insufficient number of results for individual languages. Furthermore, the heterogeneity of methodologies employed in selecting personality traits or voice parameters, the study designs, and the

[1] For a comprehensive literature review of nearly 85 years of research, please see Essay 1.

personality assessment methods challenge comparisons between studies conducted in different languages. Therefore, comparable cross-cultural studies are essential for advancing the field of vocal brand personality research, particularly given the significance of this area for companies with an international presence.

This study responds to two of the most significant limitations of applying previous research on personality perception in international marketing. It does so by investigating the vocal perception of brand personality traits in an American English-speaking context, given the prominence of voice technology in the United States. Further, this study is situated within a cross-cultural context, as it extends on the findings of the preceding German study of Essay 2.

Thus, this study is aligned with the German approach to ensure consistency in methodological and analytical procedures, facilitating a meaningful comparison of the findings. The preceding research initially identified the brand personality dimensions perceived through voice: sincerity, sensitivity, and excitement. The German BVP-Scale can assess these brand voice personality dimensions. Therefore, the **initial research objective** of the present study is to determine which brand personalities can be perceived through voice in an American English-speaking context. Consequently, an American English version of the BVP-Scale will be developed to assess the American brand voice personality dimensions.

Moreover, the German study developed a BVP-Model by examining which linear combinations of objectively measurable voice parameters induce perceptions of the identified brand voice personality dimensions. In this regard, a distinction was made between female and male voices, as the speaker's gender was found to influence personality perceptions. Consequently, **the second objective** of the present study is to develop an American BVP-Model by identifying the linear combinations of female and male voice parameters that induce the perception of American brand voice personality dimensions.

Building on the previous research objectives, the **third research objective** of this cross-cultural replication study is to determine what differences exist in vocal brand personality perceptions between the United States and Germany. Therefore, the German and American BVP-Scales are compared to see which brand personalities are perceived through voice in both cultures. Furthermore, the developed German and American BVP-Models are compared by examining the voice profiles for each brand voice personality dimension in both cultures.

Given the need for more research on cross-cultural vocal personality perception and that the United States and Germany exhibit strong similarities in their communication patterns (both are low-context and linear-active cultures), no specific hypotheses on expected differences are formulated. Consequently, as in the German study, the present study employs an exploratory approach. **Figure 2** presents the research framework for developing and comparing an American BVP-Scale and BVP-Model based on the preceding German survey.

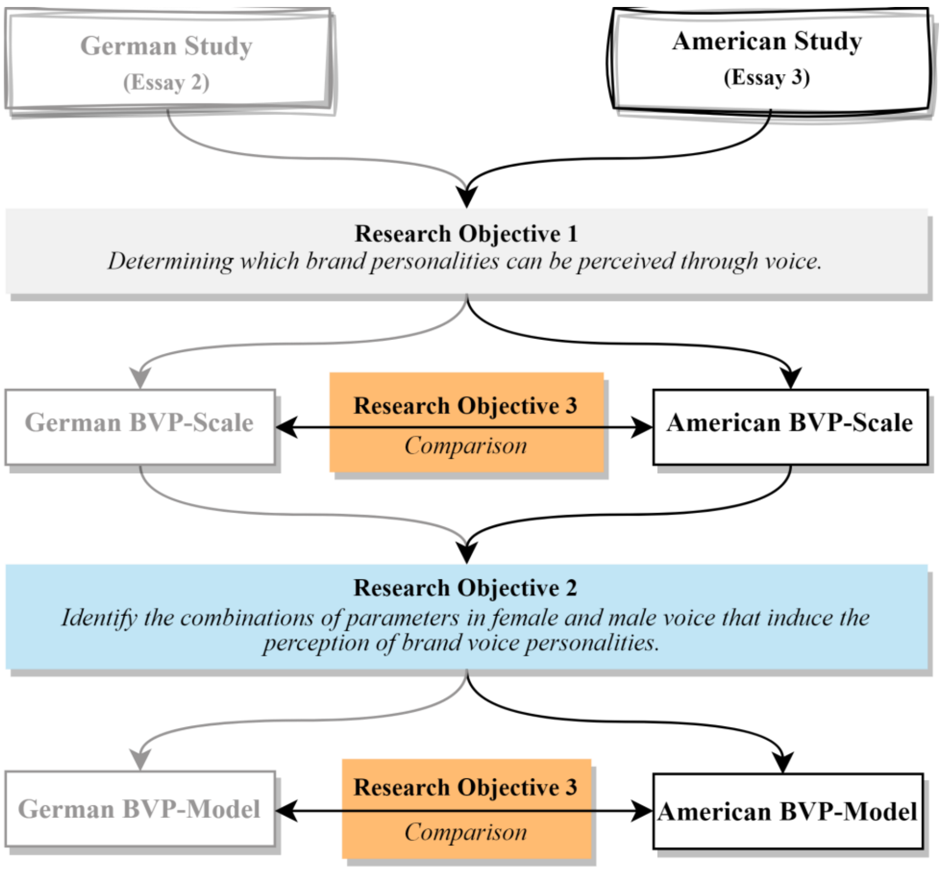


Figure 2. Research Framework of American English Perceptual Study

The scope of the present study aligns with that of the German research, which involved a correlation study with external judgments of brand personality perception through voice. This methodology consists of the evaluation of voices without prior manipulation of their acoustic measures with naïve listeners who do not regularly deal with the assessment of voices and speech, e.g., speech therapists, linguists, or voice actors (Brown et al., 1985; Koutsombogera et al., 2020; Riding et al., 2006). Unmanipulated voices allow for a natural diversity in the expression of vocal parameters. Additionally, users of voice-based technology who engage in conversations with voice assistants are likely to be unfamiliar with vocal personality judgments, which is why the experimental setup is the closest approximation to reality.

Finally, this study considers only those voice parameters that can be objectively measured and their interaction effect, i.e., multivariate analysis of voice parameters, as applied in the preceding survey. **Table 1** overviews the 15 voice parameters primarily researched and of central interest for this study, categorized according to their listener's perception and the four soundwave dimensions: timing, amplitude, frequency, and voice quality. Please see Appendix A for a detailed explanation and definition of each voice parameter and its acoustical measure.

Table 1. Voice Parameters of Interest for American Personality Perception Study

Soundwave Dimension	Acoustical Measure (Metric)	Listener's Perception
Timing	Speaking Rate (syl/s) Average Silent Pause Duration (s) Silent Pause Frequency per Minute (n/m)	Fluency of Speech
	Articulation Rate (syl/s)	Velocity of Speech
Amplitude	Intensity Variability (SD; dB)	Loudness Variability (Intonation)
Frequency	Fundamental Frequency Mean (f0; Hz)	Pitch
	Fundamental Frequency Standard Deviation (SD f0; Hz) Fundamental Frequency Range (f0-max – f0-min; st)	Pitch Variability (Intonation)
Voice Quality	h1-h2 (dB)	Creakiness; Breathiness
	Spectral Slope Spectral Tilt	Creakiness; Breathiness, Loudness (Brightness)
	Smoothed Cepstral Peak Prominence (CPPS; dB)	Breathiness
	Harmonics-to-Noise Ratio (HNR; dB) Jitter (%) Shimmer (%)	Roughness/ Hoarseness

Notes: syl = syllable; s = second; m = minute; n = number; SD = standard deviation; Hz = Hertz; st = semitones; dB = Decibel.

4. Methodology

4.1. Participants and Procedure

In this study, naïve listeners evaluated voices according to perceived brand personality traits via an online survey divided into three sections. The first section consisted of demographic and filter questions. Participants who were not native American English speakers or had an uncompensated hearing impairment were excluded. In the second section, participants listened to and rated an initial voice on a five-point scale (from 1 = ”does not apply at all” to 5 = “applies completely”) on the extent to which they associated the voice with the personality traits provided. Subsequently, the same procedure was repeated for a second voice. The voices were presented in a loop until participants had completed the respective personality ratings. Finally, in the third section, participants were asked to indicate the device they used to complete the questionnaire (speakers or headphones) and the background noise level they experienced while filling out the questionnaire. All participants provided informed consent before participating.

The personality perception rating was based on 62 personality traits derived from the brand personality scales (BPS) of Aaker (1997), Geuens, Weijters, and de Wulf (2009), and Grohmann (2009), which were also used in the German study. These scales are generalized and applicable to various contexts, such as diverse product categories. Furthermore, they apply to diverse cultural settings, including the American English culture. In total, the three BPS resulted in 66 traits. However, since four traits (aggressive, daring, down-to-earth, and sentimental) occurred twice within the three scales, only 62 traits were used in the survey. **Table 2** provides an overview of all three BPS’s personality dimensions and traits.

Table 2. Overview of Used Brand Personality Traits in Survey

Authors	BPS Dimension	BPS Traits
Aaker (1997)	Competence	confident, corporate, hard-working, intelligent, leader, reliable, secure, successful, technical
	Excitement	contemporary, cool, **daring**, exciting, imaginative, independent, spirited, trendy, unique, up-to-date, young
	Ruggedness	masculine, outdoorsy, rugged, tough, western
	Sincerity	cheerful, **down-to-earth**, family-oriented, friendly, honest, original, real, **sentimental**, sincere, small-town, wholesome
	Sophistication	charming, feminine, glamorous, good-looking, smooth, upper-class
Geuens et al. (2009)	Activity	active, dynamic, innovative
	Aggressiveness	**aggressive**, bold
	Emotionality	romantic, **sentimental**
	Responsibility	**down-to-earth**, responsible, stable
	Simplicity	ordinary, simple
Grohmann (2009)	Feminity	expresses tender feelings, fragile, graceful, sensitive, sweet, tender
	Masculinity	adventurous, **aggressive**, brave, **daring**, dominant, sturdy

Notes: The eight bolded and colored words are the four traits that occur twice within the three scales. BPS = brand personality scale.

In total, 1,138 participants living in America completed the questionnaire through an online-access panel provider. Before analyzing the data, participants who made insufficient effort to answer the questionnaire were removed (Huang et al., 2012; Leiner, 2019). Thus, 44 participants with missing answers, 53 participants who did not pass attention checks, 16 participants who have demonstrably not played the voices, and 46 participants who showed low variance in their responses in the personality trait ratings primarily due to response pattern (variance $\leq .2$ was used as the threshold) were removed (Jandura, 2018; Leiner, 2019). After this data cleaning, the remaining data from 979 participants were split according to first and second voice ratings, resulting in 1,958 individual voice ratings ($M_{age} = 34.35$; 65.6% female).

4.2. Vocal Stimulus Set

The voices used as vocal stimuli were obtained from the Buckeye corpus of conversational speech (Pitt et al., 2007). This corpus is a database of voice recordings from 40 unfamiliar speakers, both male and female, consisting of interviews on daily topics (e.g., parenting, sports, traffic) in American English. The Buckeye corpus was identified through a comprehensive screening of several repositories of databases designed for the collection of spoken corpora. The corpora and databases examined included Clarin, TalkBank, Open SLR, VoxForge, The Spoken Wikipedia Corpora, The Michigan Corpus of Academic Spoken English (MICASE), Corpora Collection Leipzig, the Hamburg Centre for Speech Corpora (HZSK) of the University of Hamburg, and the International Corpus of English (ICE) of the University of Zurich.[2] This corpus was primarily chosen because other available American English spoken language corpora were designed outside of controlled conditions, making it challenging to compare the acoustical measures. Further, the Buckeye corpus consisted of free conversations with individuals through an interview, i.e., spontaneous speech, which best matched the research purpose. Previous research indicates that personality traits are most clearly manifested in spontaneous, non-restrained speech and even more so in dialogue and interactive behavior (Johnstone & Scherer, 2000; Scherer, 1972).

The Buckeye corpus was designed to show sufficient interspeaker variation and represent a speech community. To achieve this goal, it was controlled for class and ethnicity, while the sample was limited to "middle-class Caucasians" (Kiesling et al., 2006, p. 3). Further, the sample was proportional for gender and age, with 20 speakers (10 female/10 male) in the younger age group, called "under 30", and 20 speakers (10 female/10 male) in the older age group, called "over 40".

As the corpus was funded and created by the Ohio State University, only speakers who were natives of Central Ohio were recruited. Within the United States, three broad dialect categories can be distinguished: North, South, and a "Third Dialect" (Labov, 1991, p. 30). The

[2] Clarin: clarin.eu/; TalkBank: talkbank.org/; Open SLR: openslr.org/; VoxForge: voxforge.org/; The Spoken Wikipedia Corpora: nats.gitlab.io/swc/; MICASE: quod.lib.umich.edu/; Corpora Collection Leipzig: corpora.uni-leipzig.de/en; HZSK: slm.uni-hamburg.de/hzsk/; ICE: ice-corpora.uzh.ch/

"Third Dialect" region encompasses New England, the Midland, and the West. This region is the most likely to be perceived using General American English, which refers to an accent of American English spoken by most Americans (Clopper et al., 2006). Consequently, as a state in the North Midland region, the speakers of the Buckeye corpus are likely to be perceived by Americans as speaking with a General American English dialect due to their native status as Ohio residents. This limited the bias of different regional American dialects and accents on personality perceptions in this study and enabled using the corpus (Krauss et al., 2002).

The interviews conducted for corpus creation lasted from 30 to 60 minutes. For this study, samples were cut out of each interview, which contained neutral content without emotions or controversial topics and had no interruptions through the interviewee, background noises, or linguistic breaks like, for instance, coughing, tongue clicking, or laughing. The final voice samples used as vocal stimuli were, on average, 17.05 seconds long (range: 11.4 – 20.6s), which was sufficient for personality perceptions since it has been shown that personality judgments are made unconsciously after only a few seconds in former perceptual studies (Aronovitch, 1976; McAleer & Belin, 2019; Ray, 1986). All stimuli were in mono and were Root Mean Square (RMS) normalized to 60dB using the modify command in PRAAT (version 6.4.12), an open-source linguistic software tool for phonetic and speech analysis (Boersma & Weenink, 1992-2022).

Each voice sample was rated an average of 48.95 times by participants. For an overview of the selected recordings with speaker characteristics from the Buckeye corpus for this study, please refer to Appendix B. Please refer to Appendix C for access to the study's vocal stimulus set.

4.3. Acoustical Measures

The software PRAAT (version 6.4.12; Boersma & Weenink, 1992-2022) was employed to extract the 15 acoustic measures from the 40 stimuli voices. A PRAAT script, designed by the author, was used to calculate all acoustical measures and to store their numerical output. The Prosogram plugin (version 3.00f; Mertens, 2022) was employed within the script to extract data on articulation and speaking rate, silent pause duration, mean, standard deviation, and range of fundamental frequency (f0). The Voiceprofile plugin (version 2.3.1; Mayer, 2019-2021) was employed to extract data on all spectral features, including h1-h2, spectral slope and tilt, HNR, CPPS, jitter, and shimmer. PRAAT's analysis functions extracted data on the intensity variability and silent pause frequency per minute. A silent pause was defined as a pause that lasted at least .5 seconds without any audible sounds. The ranges and mean values of the acoustical measures of the vocal stimuli set per gender are summarized in **Table 3**.

Table 3. Ranges and Means of Acoustical Measures of Vocal Stimuli Set "Buckeye"

	Female Voices		Male Voices	
Acoustical Measure (Metric)	**Range**	**Mean**	**Range**	**Mean**
Speaking Rate (syl/s)	3.29 – 5.00	4.02	1.79 – 4.49	3.38
Average Silent Pause Duration (s)	.00 - .97	.57	.00 – 1.35	.73
Silent Pause Frequency per Minute (n/m)	0 – 15	6.97	0 - 25	10.51
Articulation Rate (syl/s)	3.80 – 5.95	5.04	3.70 – 6.39	5.02
Intensity Variability (SD; dB)	7.03 – 13.01	10.69	7.67 – 15.82	11.75
Fundamental Frequency Mean (f0; Hz)	115 – 226	175.55	75 - 170	109.9
Fundamental Frequency Standard Deviation (SD f0; Hz)	15.78 – 86.09	37.99	6.95 – 44.56	19.00
Fundamental Frequency Range (f0-max – f0-min; st)	6.4 – 28.1	16.94	6.0 – 26.4	12.05
h1-h2 (dB)	(-3.45) – 4.15	(-.06)	(-3.07) – 5.91	.71
Spectral Slope	(-23.86) - (-15.62)	(-19.54)	(-27.66) - (-16.75)	(-21.71)
Spectral Tilt	(-14.41) - (-13.06)	(-13.54)	(-14.32) - (-11.82)	(-13.05)
Harmonics-to-Noise Ratio (HNR; dB)	8.56 – 17.92	14.70	7.64 – 15.69	12.10
Smoothed Cepstral Peak Prominence (CPPS; dB)	10.13 – 14.65	12.40	6.65 – 14.63	11.70
Jitter (%)	1.42 – 3.34	2.18	1.49 – 6.00	2.97
Shimmer (%)	6.14 – 11.41	7.99	7.08 – 15.77	9.36

Notes: syl = syllable; s = second; m = minute; n = number; SD = standard deviation; Hz = Hertz; st = semitones; dB = Decibel.

5. Results

The data from 979 participants (65.6% female) is utilized for statistical analyses. For a comprehensive overview of the socio-demographic and survey-related sample characteristics, please refer to Appendix D. The statistical analyses are conducted using the software R (version 4.4.1) with the packages psych (version 2.4.6.26) for exploratory factor analysis (Revelle, 2024), and lavaan (version 0.6-18) for confirmatory factor analysis and structural equation model computation (Rosseel, 2012).

5.1. Development of the American Brand Voice Personality Scale

Following the first research objective, a BVP-Scale for American English is developed to indicate which brand personality traits and dimensions can be perceived through voice alone. The item pool for the scale development consisted of 62 items for brand personality assessment. The development and evaluation of the scale involved the steps of 1) assessment of the data quality, 2) verification of data factorability, and the conduction of a series of 3) exploratory factor analyses (EFA) and 4) confirmatory factor analyses (CFA), which are described in the following sections. For the EFA and CFA, the sample is split into two equal sub-samples.

For control purposes, before scale development, whether the three established scales from which the items were drawn could be confirmed within the data set was determined. The CFAs showed that none of the BPS could be reproduced, like in the German study. Because none of the BPS met the Fornell-Larcker criterion[3] and model fit indices indicated insufficient fit, it would have been necessary to exclude several items or dimensions to achieve sufficient discriminant validity and an adequate model fit (see CFA results for all three BPS in Appendices E-G; Fornell & Larcker, 1981; Hu & Bentler, 1999).

[3] Fornell-Larcker criterion: Average Variance Extracted (AVE) > squared correlation of the latent construct with the discriminant construct

5.1.1. Assessment of Data Quality and Verification of Data Factorability

The initial step in developing the American BVP-Scale is to assess the data quality. To this end, a preliminary investigation is conducted to ascertain whether there are any semantically illogical correlations between brand personality traits. As was the case in the German study, it is anticipated that the ratings of the traits “masculine” and “feminine” are limited to the gender of the voice, which is not a personality trait. Indeed, the voices of female speakers are rated high on “feminine” and low on “masculine”, while the voices of male speakers are rated high on “masculine” and low on “feminine” (see **Table 4**). Moreover, no correlations are observed between these two personality traits and the remaining traits, except for a negative correlation of r = -.71 between them. Notably, this correlation was also the highest pairwise correlation observed in the entire data set.

An additional exploratory factor analysis of the personality traits, including the two gender-related traits “masculine” and “feminine”, reveals that these two items form a separate dimension with opposite factor loadings ($\lambda_{masculine} = -.8$; $\lambda_{feminine} = .75$). This finding confirms that there are no correlations with other items and that these items are used only to assess the speaker’s gender. To focus on brand-related personality traits, and given that the results would not add value, the items “masculine” and “feminine” are excluded from further analyses. This process yields a refined item pool of 60 brand personality traits.

Table 4. Assessment of Female and Male Speakers on Feminine and Masculine Traits

	Means of “feminine” Assessment	Means of “masculine” Assessment
Female Speakers	3.96	1.43
Male Speakers	1.40	3.69

Notes: N = 1,958 individual voice ratings.

To develop a scale from a large data set, the sample of 1,958 individual voice ratings is divided into two equivalent subsamples ($n_{subsample1} = 979$; $n_{subsample2} = 979$) according to the Solomon method, a sample splitting technique that was developed explicitly for factor analyses.

The split is successful, as the similarity is assessed through a communality ratio (S)[4] of 1, which demonstrates that both subsamples have the same amount of shared variance, as determined by the Kaiser-Meyer-Olkin (KMO) statistic ($KMO_{subsample1}$ = .97; $KMO_{subsample2}$ = .97). (Lorenzo-Seva, 2021)

To ascertain the dimensionality of the BVP-Scale for American English, an EFA is conducted using subsample 1 (n = 979 individual voice ratings). The suitability of the data for factor analysis is assessed using the KMO criterion together with the personality trait's Measure of Sampling Adequacy (MSA), Bartlett's Test of Sphericity, and the anti-image covariance matrix (Bartlett, 1950; Hair et al., 2013; Kaiser, 1974; Kaiser & Rice, 1974). All testing assumptions are met, indicating that the data set is suitable for factor analysis. This suitability is supported by the KMO measure of subsample 1 along with the MSA values of each personality trait exceeding the threshold of .5 ($KMO_{subsample1}$ = .97; lowest $MSA_{fragile}$ = .89). Additionally, Bartlett's Test of Sphericity yields a statistically significant result ($\chi2(1770) = 34{,}349.703$, $p < .001$), which tests the null hypothesis that the correlation matrix shows unrelated variables (Dziuban & Shirkey, 1974). Moreover, the anti-image covariance matrix indicates that less than 25% of off-diagonal values exceed .09, suggesting that the personality traits exhibit a high common variance and that correlations are expected (Backhaus et al., 2021).

5.1.2. Exploratory Factor Analysis

In the EFA, principal axis factoring (PAF) with oblique rotation (oblimin) with Kaiser normalization is selected. The PAF method is employed for its ability to extract latent constructs that account for the relationships among measured variables, which is in contrast to principal component analysis (PCA) used to reduce the number of variables (Backhaus et al., 2021).

[4] $S = \frac{\min(KMO1,KMO2)}{\max(KMO1,KMO2)}$. KMO = Kaiser-Meyer-Olkin statistic

Additionally, it is assumed that the total amount of variance cannot be explained and that variables are linear combinations of latent constructs and residuals (Backhaus et al., 2021). As PCA assumes that variables are linear combinations of latent constructs without residuals, attempting to explain the total variance, it would not be an appropriate extraction method. An oblique rotation is chosen because correlations between constructs in personality research are assumed (in contrast to an orthogonal rotation), and the oblimin rotation is selected due to the best model fit (DeVellis, 2017).

The scree plot test without factor rotation suggests a 6-factor model based on the Kaiser criterion (eigenvalues greater than 1.0; see Appendix H; Cattell, 1966; Zwick & Velicer, 1986). Analysis of the oblique rotated factors suggests a 5-factor model with items having factor loadings greater than or equal to |.5| (Hair et al., 2020). Following the initial rotation, the model is iteratively refined by removing items until a satisfactory model is achieved. This iterative process eliminates 39 items with factor loadings lower than |.5| and communalities lower than .4, indicating insufficient reliability and internal consistency of the model (see **Table 5**; Hair et al., 2013).

Table 5. Iterative Steps in Exploratory Factor Analysis Model Purification

	Oblique Model Rotation	Items with Factor Loadings < |.5|	Items with Communality < .4
Initial Extraction and Rotation	5-factor model with 60 items	**35** (confident, hard-working, intelligent, leader, secure, successful, contemporary, daring, imaginative, independent, spirited, unique, up-to-date, western, cheerful, family-oriented, original, small-town, wholesome, charming, glamorous, good-looking, smooth, active, dynamic, innovative, aggressive, bold, romantic, ordinary, simple, graceful, brave, dominant, sturdy)	**2** (technical, young)
1st Iteration of Purification	5-factor model with 23 items	**1** (sweet)	-
2nd Iteration of Purification	5-factor model with 22 items	**1** (outdoorsy)	-
Final Model	5-factor model with 21 items	-	-

Notes: The items excluded in the respective step are shown in brackets. **Method**: Principal Axis Factoring (PAF). **Rotation**: Oblimin with Kaiser Normalization. n = 979 individual voice ratings.

The final EFA results indicate that a 5-factor model comprising 21 items best fits the American BVP-Scale, as it shows factor loadings greater than |.5|, no cross-loadings, and communalities greater than .4 for all constructs. Further, the model shows high reliability (Cronbach's alpha > .7; item-total correlations ≥ .4) and good model fit (Standardized Root Mean Square Residual (SRMR) = .02; Root Mean Square Error of Approximation (RMSEA) = .04; Tucker-Lewis Index (TLI) = .97). Please refer to Appendix I for the final EFA result.

5.1.3. Confirmatory Factor Analysis

To validate the BVP-Scale in American English, a series of CFAs are conducted with the second subsample (n = 979 individual voice ratings). The assessment of univariate and multivariate normality indicates that the data does not follow a normal distribution (Weiber & Mühlhaus, 2014; see Appendix J). Therefore, maximum likelihood estimation with robust standard errors and the Satorra-Bentler (S-B) correction is used to account for the non-normality of the data (Hu & Bentler, 1999; Tian et al., 2001). In addition, several tests are conducted to reduce measurement error and develop a scale that measures what it is supposed to measure (validity) and is consistent in measurement (reliability). These validity and reliability tests are conducted at the indicator, construct, and model levels (see **Table 6**).

Table 6. Overview of Conducted Reliability and Validity Tests

	Test Indices & Criterion	Definition	Threshold
Indicator Level	**Reliability/ Internal Consistency**		
	Factor Loadings	Correlation of an item with a construct (Backhaus et al., 2015).	≥ |.7|
	Item-Total Correlation	Correlation of each item with the average of all items in the scale (Hair et al., 2013).	≥ .4
	Cronbach's Alpha	Measure how well a set of items assesses a single construct (Cronbach, 1951).	> .7
Construct Level	**Reliability/ Internal Consistency**		
	Composite Reliability (CR)[5]	Like Cronbach's alpha, CR measures how well a set of items assesses a single construct. It considers the actual factor loadings of each item, potentially providing a more accurate estimate of reliability (Hair et al., 2013).	> .7
	Convergent Validity		
	Average Variance Extracted (AVE)[6]	The extent to which items of a construct share a high proportion of variance in common (Backhaus et al., 2015; Hair et al., 2013).	> .5
	Discriminant Validity		
	Fornell/Larcker Criterion	The degree to which a construct is distinct from other constructs (Fornell & Larcker, 1981).	AVE > squared correlation of the latent construct with the discriminant construct
Model Level	**Validity/ Global Model Fit**		
	Satorra-Bentler Chi-square test (S-Bχ^2)	Test to assess the fit between the observed covariance matrix and the estimated covariance matrix (Backhaus et al., 2015).	p > .05
	Root Mean Square Error of Approximation (RMSEA)	**Absolute Fit Measures:** Assessment of how well the theoretical model fits the sample data (Hair et al., 2013; Hu & Bentler, 1999).	≤ .05
	Standardized Root Mean Residual (SRMR)		≤ .1
	Goodness-of-Fit (GFI)		> .9
	Normed Fit Index (NFI)	**Incremental/Comparative Fit Measures:** Assessment of the proportion improvement in fit by comparing the model of interest with a baseline model (null model; Hu & Bentler, 1999; Hair et al., 2013).	> .9
	Tucker-Lewis Index (TLI)		
	Comparative Fit Index (CFI)		

5 $Composite\ Reliability\ (CR) = \frac{(\sum \text{Standardized Factor Loadings})^2}{(\sum \text{Standardized Factor Loadings})^2 + \sum Measurement\ Error\ Variances}$

6 $Average\ Variance\ Extracted\ (AVE) = \frac{\sum (\text{Standardized Factor Loadings})^2}{\sum (\text{Standardized Factor Loadings})^2 + \sum Measurement\ Error\ Variances}$

In the initial CFA, six items (real, down-to-earth, stable, fragile, corporate, rugged) indicate low indicator reliability, as reflected by factor loadings less than |.7|. Consequently, these items are excluded from further analysis. Eliminating these items transforms the two-item factor 4 (corporate and upper class) and factor 5 (rugged and tough) into single-item factors (see Appendix I). To ensure sufficient reliability and validity, factors should comprise at least two to three items (Yong & Pearce, 2013; Zwick & Velicer, 1986). Therefore, factors 4 and 5 are eliminated, along with the remaining two items, "upper class" and "tough".

The second CFA demonstrates high reliability and validity at both the indicator and construct levels. Composite Reliability (CR) values exceed .7, AVE values exceed .5, and the model satisfies the Fornell-Larcker criterion. Nevertheless, the model needs more validity at the model level. Consequently, items are removed on an iterative basis according to their modification index (MI), which indicates the expected reduction in the Chi-square statistic when the item is removed, thereby improving the model (Bagozzi & Yi, 1988). First, the item "reliable" (MI = 91.36) and then "friendly" (MI = 43.27) are removed until the model fit is optimal based on the aforementioned fit indices. Consequently, ten items are eliminated in the described iterative process of four additional CFAs until sufficient validity and reliability are achieved and the model shows a good fit.

A model's overall construct validity is achieved when it reliably measures what it is intended to measure, with high internal consistency at both the indicator and construct levels and demonstrated convergent and discriminant validity (Hair et al., 2013). Based on the specified threshold values, the model shows a high level of construct validity subsequent to the series of six CFAs.

Thus, the final CFA results in a reflective first-order model with three brand voice personality dimensions (constructs) and eleven brand voice personality traits (items; see **Table 7**). The dimensions are labeled and defined as follows:

1) **Sincerity**: This dimension describes a brand voice personality perceived as honest, responsible, and sincere.
2) **Sensitivity**: This dimension describes a brand voice personality that is perceived as sensitive, tender, sentimental, and expresses tender feelings.
3) **Excitement**: This dimension describes a brand voice personality perceived as cool, adventurous, trendy, and exciting.

Table 7. Final Confirmatory Factor Analysis for American Brand Voice Personality Scale

Construct/ Item	Standardized Factor Loading (≥|.7|)	Item-Total Correlation (≥.4)	CR (>.7)	AVE (>.5)	Fornell/Lacker Criterion
Sincerity (α=.80)					
sincere	.78***	.72	.80	.57	Yes
honest	.76***	.74			
responsible	.71***	.68			
Sensitivity (α=.84)					
expresses tender feelings	.80***	.77	.84	.57	Yes
sensitive	.75***	.74			
tender	.75***	.75			
sentimental	.71***	.67			
Excitement (α=.84)					
exciting	.79***	.77	.84	.57	Yes
cool	.77***	.73			
trendy	.74***	.73			
adventurous	.72***	.71			

Notes: Method: Maximum likelihood estimation with robust standard errors and Satorra-Bentler (S-B) correction. n = 979 individual voice ratings; ***p < .001. Fornell/Lacker criterion: AVE > squared correlations between the constructs.
Global Model Fit Indices: S-Bχ^2 (41) = 134.89 (p < .001); χ^2/df = 3.3; Standardized Root Mean Square Residual (SRMR) = .032; Root Mean Square Error of Approximation (RMSEA) = .048; Comparative Fit Index (CFI) = .98; Tucker-Lewis Index (TLI) = .97; Goodness-of-Fit Index (GFI) = .98; Adjusted Goodness of Fit Index (AGFI) = .96; Normed Fit Index (NFI) = .97.
Abbreviations: CR = Composite Reliability; AVE = Average Variance Extracted.

Once the optimal model is identified, an investigation is conducted to ascertain whether an alternative model achieves a more favorable BVP-Scale for American English due to a distinct dimensionality (Brakus et al., 2009; Tian et al., 2001). Therefore, the following models are compared: a *null model*, assuming no correlations between items; a *1-factor model*, considering all items loaded on a single construct; a *2-factor model,* considering the dimensions sincerity and sensitivity as one construct as they show the highest correlation of .71; and a *3-factor model*, considering the constructs as proposed by the final CFA. The results support the 3-factor model with the brand voice personality dimensions sincerity, sensitivity, and excitement (see **Table 8**).

Table 8. Model Fit Indices for Competing Measurement Models

Competing Models	S-Bχ^2	df	CFI	TLI	RMSEA	SRMR	AIC
Null Model	4,114.57	55	NA	NA	NA	NA	NA
One-factor Model	1,122.77	44	.77	.72	.16	.092	31,795.832
Two-factor Model	453.37	43	.91	.89	.10	.054	31,128.430
Three-factor Model	**134.89**	**41**	**.98**	**.97**	**.048**	**.032**	**30,813.952**

Notes: The results highlighted in bold are considered the best.
Abbreviations: AIC = Akaike Information Criterion; CFI = Comparative Fit Index; df = degrees of freedom; NA = not applicable; RMSEA = Root Mean Square Error of Approximation; S-Bχ^2 = Satorra-Bentler Chi-squared; SRMR = Standardized Root Mean Square Residual; TLI = Tucker-Lewis Index.

5.2. Development of the American Brand Voice Personality Model

In accordance with the second research objective, an American BVP-Model is developed by investigating which linear combinations of voice parameters induce the perceptions of brand personalities. In doing so, a distinction is made between the perception of male and female voices. Therefore, the relationship between the 15 acoustical measures (exogenous latent variables; see **Table 1**) and the three identified brand voice personality dimensions of the American BVP-Scale, sincerity, sensitivity, and excitement (endogenous latent variables), is investigated using a structural equation model (SEM). The participants' socio-demographic data (gender, age, profession, education, income) and study-related data (background noise, device used) are used as covariates.

5.2.1. Preparatory Calculations for Structural Equation Modeling

Before conducting an SEM, it is essential to address the issue of multicollinearity between exogenous latent variables (Westlund et al., 2008). Consequently, the initial step is to examine the pairwise correlation coefficients between the 15 acoustic measures (Hair et al., 2020; Sarstedt & Mooi, 2019; see Appendix K). The pairwise correlation coefficients between f0 SD and f0 range exhibit a value of .85, indicating a strong correlation. This observation suggests the potential for multicollinearity issues. In addition, the correlation between f0 SD and f0 mean is found to have a medium strength ($r = .75$). Also, HNR shows a strong correlation with shimmer ($r = -.79$) and a medium correlation with jitter ($r = -.74$). Since f0 SD and HNR show several strong and medium pairwise correlations, it is decided to exclude these two acoustical measures from subsequent analyses.

In the next step, the remaining set of 13 acoustical measures is examined with greater detail by the computation of each item's variance inflation factor (VIF). A multiple regression is therefore conducted for each acoustic measure on all other acoustic measures. Higher VIF values indicate that the variance of an acoustic measure can be explained by the other acoustic measures in the model, which is evidence of (multi)collinearity (Sarstedt et al., 2020). A critical collinearity issue will likely arise if the VIF values exceed 5, which is established as the threshold (Backhaus et al., 2021; Sarstedt & Mooi, 2019). Items with the highest VIF values are successively eliminated until multicollinearity is no longer critical. Accordingly, in the preliminary multicollinearity assessment, jitter exhibits the highest value of 6.67 and is consequently excluded. Similarly, CPPS demonstrates a VIF value of 5.88 in the subsequent multicollinearity assessment and is excluded (see Appendix L). Eliminating these two acoustical measures leads to an overall model improvement, in which the remaining eleven acoustic measures indicate acceptable VIF values.

Given the influence of gender on auditory personality perception, the data obtained from male and female voices is examined separately (McAleer et al., 2014; Trouvain et al., 2021). Therefore, a multigroup SEM with a group code approach is employed to estimate the speaker's gender-specific effects. To compare mean differences between groups, it is necessary to test for *configural* (number of factors and the pattern of factor loadings are equivalent across groups), *metric* (magnitude of factor loadings is equivalent across groups), and *scalar* (scale intercepts are equal across groups) invariance (Steenkamp & Baumgartner, 1998). The model allows co-variance between the traits "exciting" and "adventurous" of the brand voice personality dimension excitement. This adjustment is indicated by the corresponding MI of 41.88. The results of the measurement invariance tests reveal significant differences between the configural and metric models and between the metric and scalar models (see **Table 9**). This result suggests invariance constraints are violated, so invariance is not achieved. Consequently, the values of the latent means across the two genders cannot be compared directly, necessitating separate analyses for male and female voices. This lack of invariance suggests that the perception of brand personality differs between female and male voices and supports previous research on the influence of gender in vocal personality perception.

Table 9. Measurement Invariance Testing

Invariance Test	S-Bχ^2	df	RMSEA	CFI	ΔS-Bχ^2	Δdf	ΔRMSEA	ΔCFI	p
Configural Invariance	681.151	368	.031	.965					
Metric Invariance	698.004	376	.031	.964	16.853	8	.000	.001	< .05
Scalar Invariance	737.153	384	.032	.961	39.149	8	.001	.003	<.001

Notes: S-Bχ^2 = Satorra-Bentler chi-squared; df = degrees of freedom; RMSEA = Root Mean Square Error of Approximation; CFI = Comparative Fit Index; Δ = difference in respective estimates across two nested models.

5.2.2. Multigroup Structural Equation Modeling

The structural model with gender-specific effects shows a good model fit (S-Bχ^2(368) = 681.151, p < .001; RMSEA = .03; SRMR = .02; CFI = .965, TLI = .952; Hu & Bentler, 1999). **Tables 10-12** show the estimation results of the SEM for female and male voices per brand voice personality dimension of sincerity, sensitivity, and excitement, which are presented below.

For the perception of the brand voice personality sincerity, the structural model shows significant correlations for *female voices* with speaking rate ($\beta = .28$, $z = 1.69$, $p < .1$), silent pause frequency per minute ($\beta = -.03$, $z = -2.97$, $p < .05$), spectral tilt ($\beta = .57$, $z = 3.17$, $p < .05$), and shimmer ($\beta = .07$, $z = 1.95$, $p < .05$; see **Table 10A**). Furthermore, the listener's gender ($\beta = .17$, $z = 2.37$, $p < .05$) significantly correlates with sincerity perceptions in female voices, meaning specifically, when the listener is male, sincerity in female voices is perceived. Furthermore, participants who used headphones during the survey exhibit significantly stronger perceptions of sincerity in female voices ($\beta = .15$, $z = 1.66$, $p < .1$).

For the perception of the brand voice personality sincerity the structural model shows significant correlations for *male voices* with intensity variability ($\beta = -.09$, $z = -3.58$, $p < .001$), f0 range ($\beta = .03$, $z = 2.30$, $p < .05$), h1-h2 ($\beta = .07$, $z = 2.10$, $p < .05$), and spectral tilt ($\beta = -.33$, $z = -2.39$, $p < .05$; see **Table 10B**).

Table 10. Estimated Regression Coefficients of the Structural Equation Models for Brand Voice Personality Sincerity

Relationship	Path	A. Female Voices $R^2 = .071$ β	S.E.	z-value	B. Male Voices $R^2 = .062$ β	S.E.	z-value
Acoustical Measures → Sincerity	Speaking Rate → Sincerity	.28	.14	1.69*	.06	.09	.67
	Av. Pause Duration → Sincerity	.28	.19	1.23	.04	.16	.26
	Silent Pause Frequency → Sincerity	-.03	.01	-2.97**	.00	.01	.16
	Articulation Rate → Sincerity	-.22	.12	-1.58	-.01	.08	-.06
	Intensity Variability → Sincerity	.05	.04	1.04	-.09	.02	-3.58***
	f0 Mean → Sincerity	-.02	.02	-1.11	.00	.03	-.10
	f0 Range → Sincerity	.00	.01	-.36	.03	.01	2.30**
	h1-h2 → Sincerity	.01	.02	.45	.07	.03	2.10**
	Spectral Slope → Sincerity	-.02	.03	-.73	-.02	.02	-.74
	Spectral Tilt → Sincerity	.57	.15	3.17**	-.33	.13	-2.39**
	Shimmer → Sincerity	.07	.03	1.95**	.00	.03	-.02
Covariates → Sincerity	Listener's Gender[a] → Sincerity	.17	.06	2.37**	.11	.07	1.41
	Listener's Age → Sincerity	.02	.02	.92	-.04	.02	-1.39
	Listener's Education → Sincerity	.00	.02	-.16	-.02	.02	-.80
	Listener's Profession → Sincerity	.00	.02	1.03	.01	.02	.56
	Listener's Income → Sincerity	.04	.03	.93	.04	.03	1.35
	Background Noise → Sincerity	.09	.07	1.11	.15	.09	1.50
	Device Used[b] → Sincerity	.15	.08	1.66*	.09	.08	.97

Notes: Estimates represent standardized path coefficients. Significant paths have been highlighted in blue to improve orientation. The darker the color, the higher the significance level. [a] dummy coded: "0" = female, "1" = male; [b] dummy coded: "0" = speaker, "1" = headphones; *p < .1; ** p < .05; ***p < .001.
Abbreviations: av. = average; S.E. = standard errors.

For the perception of the brand voice personality sensitivity the structural model shows significant correlations for *female voices* with intensity variability (β = -.10, z = -2.27, p < .05), f0 range (β = .03, z = 2.26, p < .05), h1-h2 (β = .05, z = 2.63, p < .05), and spectral tilt (β = .53, z = 3.10, p < .05; see **Table 11A**). A significant correlation is observed between the listener's gender (β = .24, z = 3.40, p < .001) and income (β = .06, z = 2.11, p < .05) with sensitivity perceptions, indicating that males or listeners with more income exhibit a preference for sensitivity in female voices.

For the perception of the brand voice personality sensitivity, the structural model shows significant correlations for *male voices* with the speaking rate (β = -.32, z = -3.41, p < .001), intensity variability (β = -.11, z = -4.11, p < .001), and h1-h2 (β = .08, z = 2.53, p < .05; see **Table 11B**). The gender (β = .16, z = 2.24, p < .05), age (β = -.08, z = -3.47, p < .001), and income (β = .06, z = 1.92, p < .1) of the listener correlate with sensitivity perceptions, indicating that specifically when listeners are male, younger, or have more income sensitivity perceptions increase in male voices. Additionally, the two survey-related variables, background noise (β = .20, z = 2.08, p < .05) and device used (β = .17, z = 1.83, p < .1) demonstrate significant correlations. Consequently, participants who used headphones and had stronger background noise while completing the survey perceive male voices as increasingly sensitive.

Table 11. Estimated Regression Coefficients of the Structural Equation Models for Brand Voice Personality Sensitivity

		A. Female Voices $R^2 = .124$			B. Male Voices $R^2 = .098$		
Relationship	**Path**	**β**	**S.E.**	**z-value**	**β**	**S.E.**	**z-value**
Acoustical Measures → Sensitivity	Speaking Rate → Sensitivity	-.15	.13	-.97	-.32	.09	-3.41***
	Av. Pause Duration → Sensitivity	-.08	.18	-.38	.03	.15	.21
	Silent Pause Frequency → Sensitivity	.00	.01	.40	-.01	.01	-.77
	Articulation Rate → Sensitivity	.21	.11	1.64	.13	.08	1.63
	Intensity Variability → Sensitivity	-.10	.04	-2.27**	-.11	.02	-4.11***
	f0 Mean → Sensitivity	.01	.02	.40	-.01	.03	-.23
	f0 Range → Sensitivity	.03	.01	2.26**	.01	.01	.94
	h1-h2 → Sensitivity	.05	.02	2.63**	.08	.03	2.53**
	Spectral Slope → Sensitivity	.04	.03	1.20	.00	.02	-.12
	Spectral Tilt → Sensitivity	.53	.15	3.10**	-.14	.13	-1.05
	Shimmer → Sensitivity	-.02	.03	-.61	-.04	.03	-1.16
Covariates → Sensitivity	Listener's Gender[a] → Sensitivity	.24	.06	3.40***	.16	.07	2.24**
	Listener's Age → Sensitivity	-.02	.02	-.63	-.08	.02	-3.47***
	Listener's Education → Sensitivity	-.03	.02	-1.21	.00	.02	.11
	Listener's Profession → Sensitivity	.01	.01	.54	.02	.02	1.22
	Listener's Income → Sensitivity	.06	.02	2.11**	.06	.03	1.92*
	Background Noise → Sensitivity	.03	.07	.30	.20	.09	2.08**
	Device Used[b] → Sensitivity	.10	.08	1.10	.17	.09	1.83*

Notes: Estimates represent standardized path coefficients. Significant paths have been highlighted in blue to improve orientation. The darker the color, the higher the significance level. [a] dummy coded: "0" = female, "1" = male; [b] dummy coded: "0" = speaker, "1" = headphones; *p < .1; ** p < .05; ***p < .001.
Abbreviations: av. = average; S.E. = standard errors.

For the perception of the brand voice personality excitement, the structural model shows significant correlations for *female voices* with average pause duration (β = -.63, z = -2.70, p < .05), h1-h2 (β = .04, z = 1.97, p < .05), and spectral slope (β = .08, z = 2.66, p < .05; see **Table 12A**). Further significant correlations are identified between the listener's gender (β = .26, z = 3.54, p < .001), age (β = -.08, z = -3.74, p < .001), and income (β = .07, z = 2.52, p < .05) with excitement perceptions. Consequently, when the listener is male, younger, or has more income, excitement perceptions increase in female voices.

For the perception of the brand voice personality excitement, the structural model shows significant correlations for *male voices* with silent pause frequency per minute (β = -.03, z = -2.94, p < .05) and articulation rate (β = .29, z = 3.54, p < .001; see **Table 12B**). As for the female voices, when the listener is male (β = .19, z = 2.56, p < .05) or younger (β = -.09, z = -3.99, p < .001), excitement perceptions increase in male voices.

Table 12. Estimated Regression Coefficients of the Structural Equation Models for Brand Voice Personality Excitement

		A. Female Voices $R^2 = .107$			B. Male Voices $R^2 = .104$		
Relationship	**Path**	**β**	**S.E.**	**z-value**	**β**	**S.E.**	**z-value**
Acoustical Measures → Excitement	Speaking Rate → Excitement	-.01	.15	-.06	-.11	.10	-1.18
	Av. Pause Duration → Excitement	-.63	.23	-2.70**	.13	.17	.83
	Silent Pause Frequency → Excitement	-.01	.01	-1.11	-.03	.01	-2.94**
	Articulation Rate → Excitement	.18	.13	1.30	.29	.09	3.54***
	Intensity Variability → Excitement	-.05	.04	-1.24	-.01	.03	-.20
	f0 Mean → Excitement	.01	.02	.41	-.04	.03	-1.42
	f0 Range → Excitement	.00	.01	.33	.01	.01	.59
	h1-h2 → Excitement	.04	.02	1.97**	-.01	.03	-.36
	Spectral Slope → Excitement	.08	.03	2.66**	.04	.03	1.49
	Spectral Tilt → Excitement	.19	.17	1.14	-.13	.15	-.93
	Shimmer → Excitement	.00	.03	-.02	.00	.03	-.06
Covariates → Excitement	Listener's Gender[a] → Excitement	.26	.07	3.54***	.19	.08	2.56**
	Listener's Age → Excitement	-.08	.02	-3.74***	-.09	.02	-3.99***
	Listener's Education → Excitement	-.03	.02	-1.40	-.01	.02	-.25
	Listener's Profession → Excitement	.02	.02	.99	-.01	.02	-.56
	Listener's Income → Excitement	.07	.03	2.52**	.03	.03	.91
	Background Noise → Excitement	-.06	.09	-.66	-.02	.11	-.21
	Device Used[b] → Excitement	.07	.09	.75	.06	.09	.68

Notes: Estimates represent standardized path coefficients. Significant paths have been highlighted in blue to improve orientation. The darker the color, the higher the significance level. [a] dummy coded: "0" = female, "1" = male; [b] dummy coded: "0" = speaker, "1" = headphones; *p < .1; ** p < .05; ***p < .001.
Abbreviations: av. = average; S.E. = standard errors.

6. Discussion

6.1. American Vocal Brand Personality Perception

The developed BVP-Scale for American English consists of three distinct brand voice personality dimensions: sincerity, sensitivity, and excitement. Brand voice personality sincerity is described as an honest, responsible, and sincere brand. Sensitivity is defined as a brand personality perceived as sensitive, tender, sentimental, and expressing tender feelings. Finally, excitement describes a cool, adventurous, trendy, and exciting brand. The perceptions of these three brand voice personalities are tested with eleven acoustical measures to develop the American BVP-Model that determines how female and male voices induce the perception of brand voice personality dimensions (see Appendix M). In accordance with the American BVP-Model, the following voice profiles are derived for each brand voice personality dimension to provide an overview of which linear combinations of acoustical measures are (un)favorable for the respective personality perception in female and male voices.

6.1.1. Sincerity in Brand Voice

Brands are perceived as sincere when the *female voice* is creaky and rough and shows no brightness in terms of loudness. This profile is based on the correlations between spectral tilt and shimmer with sincerity ratings. Additionally, a fluent speaking style is related to increased sincerity perceptions which is indicated by the correlations of a fast speaking rate and a low frequency of silent pauses with sincerity assessments (see **Table 13A**). The voice of sample 16 is rated high on sincerity ($M_{sincerity}$ = 4.1; see Appendix C) and is thus illustrative of a voice suitable for representing a sincere female brand. The speaker's vocal creakiness (spectral tilt = -13.23) and roughness (shimmer = 7.92%) are audible, and the voice can be described as tranquil and not particularly loud, consistent with the significant results concerning the brightness of the voice. The speaking rate is 4.2 syllables per second, which is slightly above average fluency ($M_{speaking\_rate}$ = 4.02 syl/s within female stimuli set). Regarding the fluency of speech, the results align with prior research on the perception of reliability, a personality trait closely associated with sincerity. Charoenruk and Olson (2018) report a positive correlation between increasing speaking rates and reliability ratings in female and male voices. Finally, the listener's gender is a significant factor, with male listeners perceiving higher sincerity when the voice representing the brand is female.

The *male profile* for sincere brands indicates that the appropriate brand voice is characterized by a breathy quality and a bright speech intensity (see **Table 13B**). The significant correlations between spectral tilt and h1-h2 with sincerity ratings evidence this profile. Regarding intonation, the results are contradictory regarding the perception of sincerity for male brands at first sight. The negative correlation with intensity variability indicates that a monotone speech style benefits sincerity perception. Conversely, the positive correlation with the f0 range indicates a preferred dynamic speech style. Both intonation information is compatible, which is also observed in the two voices of sample 38 ($M_{sincerity}$ = 4.0) and sample 19 ($M_{sincerity}$ = 3.9), two of the highest-rated male voices regarding sincerity (see Appendix C). Both speakers increase the pitch of their voices to emphasize particular words or statements, resulting in a higher variability in pitch (f0 range).[7] Through these brief pitch changes, both speakers demonstrate high mean pitches (f0 $mean_{sample38}$ = 111Hz; f0 $mean_{sample19}$ = 129Hz), although their fundamental pitch can be described as deep. At the same time, the speakers show a monotonous use of their loudness during the speech, resulting in a below-average intensity variability of 9.5 and 11.00dB, respectively ($M_{intensity\_variability}$ = 11.75dB within the male stimuli set). These observations indicate that the speakers use their pitch instead of their loudness for emphasis. In conclusion, male brand voices should demonstrate dynamic pitch variability and monotonic loudness

[7] The speaker of sample 38 increases the pitch of his voice when stating "computers in the room" (4s) and "kindergarten" (9s). The speaker of sample 19 increases the pitch of his voice when saying "had" (10s) and "provide to us" (12s).

variability when speaking. Additionally, they should exhibit breathiness and brightness in their voices, as these qualities contribute to the perception of sincerity.

Table 13. American English Brand Voice Profile of Sincerity

	Voice Parameters	
	Listener's Perception	**Acoustical Measure (Metric)**
A. Female	fluent speech	increased speaking rate (syl/s) decreased silent pause frequency per minute (n/m)
	roughness	increased shimmer (%)
	creakiness no brightness	flat spectral tilt
B. Male	dynamic intonation through pitch	increased f0 range (Hz)
	monotone intonation through loudness	decreased intensity variability (dB)
	brightness	steep spectral tilt
	breathiness	increased h1-h2 (dB) steep spectral tilt

Notes: f0 = fundamental frequency; syl = syllable; s = second; m = minute; n = number; Hz = Hertz; dB = Decibel.

6.1.2. Sensitivity in Brand Voice

The *female voice* for sensitive brands is creaky, breathy, and without brightness or articulatory clarity. This profile is evidenced by the significant correlations between the acoustic measures spectral tilt and h1-h2 with sensitivity ratings. Moreover, female voices should demonstrate dynamic intonation in terms of pitch variability, whereas in terms of loudness variability, they should exhibit a monotone intonation (see **Table 14A**). As previously illustrated with the male sincerity voice profile, the speakers of the highly sensitive voices of sample 1 ($M_{sensitivity}$ = 3.8) and sample 7 ($M_{sensitivity}$ = 3.6) raise their pitch to stress words or statements, increasing their pitch variability (see Appendix C).[8] Furthermore, the speaker of sample 7 has a relatively low pitch (f0 mean) of 144Hz, as evidenced by her use of vocal fry. This creaky

[8] The speaker of sample 1 increases the pitch of her voice when stating "it didn't affect me at all" at the beginnings of her speech. The speaker of sample 7 increases the pitch of her voice when saying "little" (13s) and "school" (15s).

vocal quality occurs when speakers lower their pitch to the lowest register they can produce, typically at the end of their sentences (Anderson et al., 2014). As creakiness in a voice positively correlates with sensitivity perceptions, speakers with vocal fry will likely show a high pitch variability when selective pitch raises are made – the change between high and low pitch results in high variability. Further, both speakers show low loudness variability, which leads to the perception of a monotone speaking style in terms of intonation. This finding is consistent with the observation that the female-sensitive brand voice should be less loud in brightness and articulatory clarity, as evidenced by the spectral tilt correlation. In addition to the female voice profile described for sensitive brand perceptions, male and higher-income listeners perceive a higher level of sensitivity in female voices.

The *male brand voice profile* for sensitivity perception exhibits less complexity and diversity than the female profile. The results demonstrate that a breathy voice with monotone intonation and hesitant speech induces sensitivity perceptions in males (see **Table 14B**). The correlations between intensity variability, speaking rate, and h1-h2 with sensitivity ratings indicate this profile. The voices of sample 22 ($M_{sensitivity}$ = 3.3) and sample 28 ($M_{sensitivity}$ = 3.1) are rated the highest on sensitivity and are thus illustrative of a voice suitable for representing a sensitive male brand (see Appendix C). Both samples exhibit a calm and thoughtful manner of speaking, characterized by a monotone intonation and brief and frequent pauses (as observed in voice sample 22) or longer and less frequent pauses (as observed in voice sample 28). This pausing behavior results in a slow speaking rate of 2.4 and 1.8 syllables per second, respectively. This slow speaking rate contributes to both speakers' observed hesitant speaking style. In addition, sensitivity in male voices is more likely to be perceived by male, younger, or higher-income listeners, which should be considered when addressing a specific target audience.

Table 14. American English Brand Voice Profile of Sensitivity

	Voice Parameters	
	Listener's Perception	**Acoustical Measure (Metric)**
A. Female	dynamic intonation through pitch	increased f0 range (Hz)
	monotone intonation through loudness	decreased intensity variability (dB)
	breathiness	increased h1-h2 (dB)
	creakiness no brightness	flat spectral tilt
B. Male	hesitant/ less fluent speech	decreased speaking rate (syl/s)
	monotone intonation through loudness	decreased intensity variability (dB)
	breathiness	increased h1-h2 (dB)

Notes: f0 = fundamental frequency; syl = syllable; s = second; Hz = Hertz; dB = Decibel.

6.1.3. Excitement in Brand Voice

The perception of excitement is associated with a creaky and breathy *female voice*, which has a less bright and clear articulation and a fluent speaking style. These findings are based on the correlations between spectral slope, h1-h2, and average pause duration with excitement ratings (see **Table 15A**). The most exciting voice of sample 39 ($M_{excitement}$ = 2.9) shows a unique feature in terms of fluency (see Appendix C). The speaker does not use silent pauses during her speech,[9] which leads to an above-average speaking rate of 4.4 syllables per second ($M_{speaking\_rate}$ = 4.02 syl/s within the female stimuli set). This finding is consistent with previous research on human personality perception in a broader sense, as preference for a fluent speaking style was found to positively influence perceptions of a dynamic and active personality in women (Addington, 1968; Aronovitch, 1976). The positive impact of breathiness and creakiness on excitement perception is surprising, as these voice qualities intuitively are more connected with relatively calm personalities. This assumption would be supported by the fact that

[9] A silent pause is defined as a pause that lasted at least .5s without any audible sounds. Therefore, silent pauses shorter than .5s long can still occur in the voice samples.

for the perception of a female-sensitive brand, creakiness and breathiness are also favorable (see **Table 15A**). Nevertheless, the linear combination of the voice parameters may explain the similar voice profiles of sensitivity and excitement concerning the voice qualities. While for the sensitivity brand personality, breathiness and creakiness are combined with dynamic pitch variability and monotone loudness variability, breathiness and creakiness in excitement perceptions are combined with speech fluency. Combining the same voice qualities with other voice parameters results in varying brand personality perceptions. This crucial role of the linear combination of voice parameters must be considered when designing an exciting brand personality voice represented by a female.

The perception of excitement in *male voices* is dominated by combining two timing voice parameters: silent pause frequency per minute and articulation rate. Thus, excitement is perceived through fluent and quick speech (see **Table 15B**). Except for the impact of voice qualities, there are similarities concerning the preferred fluency in exciting speech with the female profile and with previous research on the perception of an active personality conducted by Addington (1968). Voice sample 30 ($M_{excitement} = 3.0$) represents a suitable male voice for an exciting brand (see Appendix C). Even though the speaker hesitates long at the beginning of his speech,[10] he continues to speak without pauses, with an articulation rate of 4.65 syllables per second.

Male and younger listeners are likelier to identify excitement in female and male voices. Additionally, female voices are also perceived as exciting by higher-income listeners. Notably, most of the voices that receive high ratings for excitement are from speakers of the "under 30" age group, e.g., voice samples 30, 39, 37, and 8 (see Appendix C). This observation might indicate that younger speakers may speak in a way that is perceived as being exciting,

[10] The hesitation is the result of a series of five pauses, with an approximate duration of 2 seconds (silent pause of .7s; filled pause ("eh") of .3s; silent pause of .5s; filled pause ("like") of .2s; silent pause of .3s).

particularly by listeners from the same age group. Thus, it is recommended to use younger female or male voices for exciting brands as excitement perceptions increase. Further, a young target group is further favorable for excitement perceptions.

Table 15. American English Brand Voice Profile of Excitement

	Voice Parameters	
	Listener's Perception	**Acoustical Measure (Metric)**
A. Female	fluent speech	decreased average duration of silent pauses (s)
	breathiness	increased h1-h2 (dB)
	creakiness no brightness	flat spectral slope
B. Male	fluent speech	decreased silent pause frequency per minute (n/m)
	quick speech	increased articulation rate (syl/s)

Notes: syl = syllable; s = second; m = minute; n = number; dB = Decibel.

6.2. Cross-Cultural Evaluation of Vocal Brand Personality Perception

The third research objective of this study is to determine what cultural differences exist in vocal brand personality perceptions between the United States and Germany. Therefore, based on the results of the present American study and those of the German research presented in Essay 2, the extent to which the developed BVP-Scales of both cultures resemble each other is discussed. Further, the developed BVP-Models are compared by examining the derived brand voice profiles for each brand voice personality dimension in both cultures.

6.2.1. Comparison of Brand Voice Personality Scales

The German BVP-Scale was developed using data from 2,000 German participants rating 96 voices of semi-spontaneous speech on 64 brand personality traits, which were German translations of the traits taken from the BPSs of Aaker (1997), Geuens et al. (2009), and Grohman (2009). The American study employed the same brand personality traits in English.

In contrast, the item pool comprised 62 traits due to four traits occurring twice within the English versions of the BPSs. The necessity for developing a BVP-Scale in both languages was identified because of the inability to reproduce neither of the three BPSs in both data sets. Furthermore, the initial data assessment demonstrated that the personality traits "masculine" and "feminine" were exclusively associated with the speaker's gender and exhibited no significant correlations with other brand personality traits. Consequently, these two traits were excluded from both BVP-Scales, and numerous other brand personality traits were eliminated until satisfactory models were achieved. Through a series of EFA and CFA, three-dimensional BVP-Scales were developed for German and American English (see Appendix N for the German CFA result). Both scales consist of the brand voice personality dimensions of sincerity, sensitivity, and excitement and encompass eleven brand voice personality traits (see **Table 16**).

Table 16. German and American Brand Voice Personality Scales

German Brand Voice Personality Traits	Brand Voice Personality Dimension	American English Brand Voice Personality Traits
sincere		sincere
honest	**Sincerity**	honest
reliable		responsible
sensitive		sensitive
sentimental	**Sensitivity**	sentimental
fragile		tender
smooth		expresses tender feelings
exciting		exciting
adventurous	**Excitement**	adventurous
spirited		trendy
daring		cool

Notes: Because the two brand voice personality scales exhibit identical naming and number of dimensions, the dimensions displayed in the center reflect those of both scales.

Both dimensions of sincerity consist of brand voice personality traits that belong to a brand voice's perceived sincerity and honesty. The primary distinction between the two dimensions is that "reliable" strongly correlates with sincerity and honesty in a German-speaking context. In contrast, in an American English-speaking context, “responsible” strongly correlates with sincerity and honesty traits. Reliability is connected to feelings of trustworthiness and confidence; responsibility can be associated with accountability (Bovens, 2007). Therefore, sincere German brands are perceived as trustworthy and reliable. In comparison, American sincere brands are held accountable for their actions based on the associated responsibility.

The brand voice dimensions of sensitivity of both languages are strongly comparable. Both dimensions describe emotional brands based on the brand voice personality traits “sensitive” and “sentimental”. Further, the German sensitivity dimension encompasses the traits “fragile” and “smooth”, and the American sensitivity dimension encompasses “tender” and “expresses tender feelings” as traits. The distinctions between these two traits are subtle, as “smooth” and “tender” are synonymous. Furthermore, the traits “fragile” and “expresses tender feelings” describe a sensitive brand’s emotional character.

The most significant differences between the two BVP-Scales are observed within the excitement dimensions. Both dimensions include “exciting” and “adventurous” as brand voice personality traits, describing an exciting brand with an enjoyment of experiences. However, exciting German brands are associated with “daring” and “spirit”, which is not the case for exciting American brands, which are associated with a relatively “cool” and “trendy” character. Moreover, the German translation of “spirited” (i.e., “temperamentvoll”) means being lively and temperamental. This trait contrasts with the coolness associated with an American exciting brand.

In conclusion, the three dimensions of the German and American BVP-Scales exhibit high comparability despite marginal differences in the item composition. Therefore, similar brand personality dimensions are perceived through voice in both cultures. The most remarkable difference lies in the excitement dimensions, which must be considered when comparing the two personalities and their vocal perception. The item composition of the respective dimensions may be attributed to cultural differences in the perception of personalities in general. Additionally, differences could arise from the translations of the traits, which result in slight semantic variations that significantly impact the perception of voices.

6.2.2. Comparison of Brand Voice Personality Models

The BVP-Models are developed based on the BVP-Scales for German and American English. Therefore, the relationships between the three identified brand voice personality dimensions, sincerity, sensitivity, and excitement, and the linear combination of eleven acoustical measures are examined in both cultures. All measures derived from previous perceptual studies identifying significant correlations between personality traits and the human voice are acoustically measurable. Based on the identified effects of speaker gender on the perception of brand personalities, a multi-group SEM with a group code approach is applied in both languages.

Appendices O-Q present the SEM estimation results for female and male voices per brand voice personality dimension for German and American English, which form the final BVP-Models. These results are used to derive the voice profiles for each brand voice personality dimension and as the basis for the subsequent comparison of German and American English brand voice profiles.

The German and American *female voice profiles* for the perception of the brand voice personality dimension **sincerity** are comparable (see **Table 17A**). A female voice should speak fluently in both languages, exhibit no brightness, and include creakiness. The only difference is that a German sincere brand should be breathy and creaky, whereas an American English brand should be rough and creaky. Therefore, it is possible for international acting organizations with brands that identify themselves with a sincere personality to choose a female voice with fluent speech and creakiness and adjust this voice by adding breathiness or roughness depending on the customer's culture. However, it is essential to ensure that the naturalness of the voice is not changed when voice qualities are modified. Voice qualities often describe both periodic and aperiodic aspects of the voice at the frequency or intensity level, which is why these parameters correlate with each other and with other voice parameters (Hildebrand et al., 2020; E. Keller, 2005).

The *male voice profiles* of a **sincere** brand exhibit substantial cultural differences, as illustrated in **Table 17B**. In Germany, brand sincerity is perceived through a creaky voice, which speaks fluently without brightness in terms of loudness. Conversely, sincere American brands need voices with brightness and breathiness, dynamic pitch intonation, and monotone loudness intonation. An explanation for this difference is that the sincerity of male brands is perceived differently in German than in American-English culture. Another potential explanation for the discrepancy between the two brand voice profiles is the differing item composition within the sincerity dimensions. As previously outlined, German sincere brands are perceived as reliable, while American sincere brands are perceived as responsible (see **Table 16**). This subtle difference in the associated traits of sincerity in brands can be crucial in shaping the vocal perception. Consequently, two distinct male voices should be selected for sincere brands in both cultures.

Table 17. German and American Brand Voice Personality Profiles of Sincerity

	German Brand Voice Profile	American Brand Voice Profile
A. Female	fluent speech	fluent speech
	creakiness	creakiness
	no brightness	no brightness
	breathiness	roughness
B. Male	no brightness	brightness
	creakiness	breathiness
	fluent speech	dynamic intonation through pitch
		monotone intonation through loudness

Notes: These brand voice profiles are derived based on the multigroup SEM results on the perception of sincerity through the linear combination of 11 acoustical measures (see Appendix O).

The voice profiles for **sensitive** *female brands* in German and American English share the need for breathiness and a low loudness variability, resulting in monotone intonation (see **Table 18A**). Additionally, a low pitch and a fluent and slow speaking style increase sensitivity perceptions in German voices, whereas roughness in voice decreases sensitivity perceptions and should, therefore, be avoided. Next to the breathiness and monotone loudness intonation, the female voice of a sensitive American brand needs to be creaky, not bright, and dynamic in pitch intonation. Thus, for the perception of sensitivity in American brands, voice parameters, which affect the quality of voice and intonation in speech, are crucial. In contrast, timing and frequency parameters play a decisive role in the perception of sensitivity in German female voices.

The voice profiles for **sensitive** *male brands* in German and American English share the need for breathiness (see **Table 18B**). Other than that, the two profiles differ greatly from each other. In addition to the voice's breathiness, the German sincere brand voice should be creaky, whereas roughness and brightness should be avoided as they decrease the sensitivity perceptions of male voices. In contrast, in addition to the voice's breathiness, the American English male brand voice should be monotone in loudness intonation and hesitant in speech through a decreased speaking rate so that a sensitive brand personality is perceived.

In conclusion, brand voice sensitivity perceptions differ in gender and culture. The combination of crucial voice parameters describing each sensitive brand voice is entirely different, indicating that unique brand voices are necessary to capture the influence of gender and cultural context. Despite the observed differences in female brand voice profiles between the two cultures, a commonality exists that German and American males, compared to female listeners, tend to perceive sensitivity in brands, regardless of the voice gender. This finding is especially favorable for sensitive brands whose target groups are predominantly male.

Table 18. German and American Brand Voice Personality Profiles of Sensitivity

	German Brand Voice Profile	American Brand Voice Profile
A. Female	monotone intonation through loudness	monotone intonation through loudness
	breathiness	breathiness
	fluent speech	dynamic intonation through pitch
	slow speech	creakiness
	low-pitched voice	no brightness/ articulatory clarity
	no roughness	
B. Male	breathiness	breathiness
	creakiness	hesitant/ less fluent speech
	no roughness	monotone intonation through loudness
	no brightness	

Notes: These brand voice profiles are derived based on the multigroup SEM results on the perception of sensitivity through the linear combination of 11 acoustical measures (see Appendix P).

In both cultures, *female voices* of **exciting** brands should be breathy and creaky (see **Table 19A**). Furthermore, brightness in voice reduces excitement perceptions in German and American English and should be avoided in exciting brand voices. However, while American English female brands use short silent pauses during a speech to be perceived as exciting, German female brands use long pauses, decreasing the speech's fluency. More specifically, the pauses are observed to be longer, and the speaking rate between these voice breaks is increased. This observation describes an interesting speaking style for exciting German brands in which long silent pauses are utilized, but the speech is delivered relatively quickly between such voice breaks. To conclude, the voice profiles of exciting bands in both cultures are strongly comparable but slightly differ in timing parameters. The manipulation of timing parameters is generally more straightforward than the manipulation of voice qualities, as their characteristics can be more accurately quantified (Winn, 2020). For example, the speaking rate and the length of pauses can be altered by cutting and rearranging speech segments without significantly affecting the naturalness of the voice. Therefore, international brands could choose the same female voice for Germany and the United States and adjust the timing parameters depending on the cultural preferences for excitement perception.

As noted in the evaluation of the American English *male voice* for **exciting** brands, fluent and quick speech increases excitement judgments (see **Table 19B**). In contrast, excitement in German brand voices is perceived through hesitant speech and a voice with a breathy and creaky quality. Thus, the German male brand voice profile is comparable to the female profile, whereas it can be described as the opposite of the American male profile. In Germany, similar voices can be employed to represent exciting brands. In contrast, different voices must be selected in the United States, depending on whether the speaker is male or female. At this point, it is essential to emphasize that the discrepancy between the male profiles could be a product of the different item compositions of the excitement dimensions of both cultures. German exciting brands are associated with traits like daring and spirit, whereas exciting American brands are associated with traits like cool and trendy.

In addition to the effect of voice parameters on perceptions of excitement, the findings reveal that male and younger listeners perceive higher excitement levels in both female and male voices. Thus, it is recommended that in both cultures, voices belonging to younger females or males are employed to convey the quality of excitement associated with brands as excitement perceptions increase.

Table 19. German and American Brand Voice Personality Profiles of Excitement

	German Brand Voice Profile	American Brand Voice Profile
A. Female	breathiness	breathiness
	creakiness	creakiness
	no brightness	no brightness
	hesitant speech through the use of long pauses	fluent speech through the use of short pauses
	fluent speech through speaking rate	
	no roughness	
B. Male	hesitant speech through the use of long pauses	fluent speech through a reduced number of silent pauses
	breathiness	quick speech
	creakiness	
	no roughness	
	no brightness	

Notes: These brand voice profiles are derived based on the multigroup SEM results on the perception of excitement through the linear combination of 11 acoustical measures (see Appendix Q).

In conclusion, the cross-cultural comparison of brand voice profiles reveals that the profiles exhibit more differences than similarities in most cases. The results highlight the necessity for organizations with an international presence to select or design distinct brand voices. The female voice profiles of sincere and exciting brands have the most remarkable resemblance, although minor variations in voice quality and timing parameters should be considered. The voice profiles of sensitive brands are distinct in both cultures, as are the male voice profiles of sincere brands. The male brand voice profile for exciting brands can be described as being the opposite in Germany and the United States.

Despite the comparable communication patterns of both cultures (low-context and linear-active cultures) and the perception of similar brand personality dimensions through voice, the decisive voice parameters that induce the perception of brand voice personality dimensions are different. These findings support the theory that cross-cultural variations in learned social interaction norms and rules are crucial in vocal personality perception. In line with the limited number of studies investigating perceptual differences between German and American English, this study demonstrates that an opposite use of the same voice parameters also induces the same brand personality perceptions. For instance, exciting American brands are characterized by fluent and rapid speech, whereas exciting German brands exhibit a hesitant speech pattern accompanied by breathy and creaky voice quality. These findings are comparable to those on extraversion. A slow articulation rate in female Germans was linked to increased levels of extraversion (Michalsky et al., 2020), and a fast speaking rate improved perceptions of extraversion in female Americans (Addington, 1968; Aronovitch, 1976). The resemblance between an exciting brand voice personality and an extroverted human personality may be attributed to the fact that one facet of the Big Five extraversion dimension is "excitement seeking" (McCrae et al., 2005).

7. Conclusion

In the context of voice AI, international acting companies must demonstrate an understanding of the significant impact of culture as an acoustic cue and consider adapting their brand's voice to align with the cultural context in which they seek to interact with customers. The representation of a brand through its voice affects the perception of the brand personality, which in turn influences brand identification (Nam et al., 2011). This perception is shaped not only by the voice and its acoustic characteristics but also by the cultural context of the customer (Krauss et al., 2002). Despite this crucial role culture plays in shaping perceptions of personality and the significance of cross-cultural comparisons in voice marketing, research in this domain, particularly concerning the vocal perception of brand personalities, is missing. This study addressed this gap by conducting the first cross-cultural analysis of vocal brand personality perception in Germany and the United States.

The present study initially identified the brand personality dimensions that can be perceived through voice alone in an American English-speaking context. The American BVP-Scale, designed to assess a brand's voice personality, comprises three distinct dimensions: sincerity, sensitivity, and excitement. Further, an American BVP-Model was developed by identifying the linear combinations of female and male voice parameters that induce the perception of the identified brand voice personality dimensions. The results were used to derive voice profiles for each brand voice personality dimension.

Sincerity in American brands is perceived when *female voices* are fluent in speaking, creaky, rough, and show no brightness; and *male voices* are breathy, bright, and demonstrate dynamic pitch variability and monotonic loudness variability when speaking. **Sensitivity** in American brands is perceived when *female voices* are creaky, breathy, and without brightness but with a dynamic pitch intonation and a monotone loudness intonation; and *male voices* are breathy with monotone intonation and hesitant speech. **Excitement** in American brands is perceived when *female voices* are breathy and creaky, have a less bright and clear articulation and a fluent speaking style; and *male voices* demonstrate a fluent and quick speaking style.

In a final step, the American and German BVP-Scales and BVP-Models were compared to identify cross-cultural differences in the perception of brand personalities. The present study aligned with the German methodological and analytical procedures, facilitating a meaningful comparison of the findings. Both scales were found to consist of the brand voice personality dimensions of sincerity, sensitivity, and excitement and encompass eleven brand voice personality traits. While some minor differences in the item composition were noted, it was concluded that both cultures perceive similar brand personality dimensions through voice.

In contrast, comparing the American and German BVP-Models and their derived brand voice profiles revealed that the perception of the brand voice personality dimensions is induced through different combinations of voice parameters. The female brand voice profiles of brands perceived as sincere and exciting have the most remarkable resemblance to one another. The voice profiles of sensitive brands are distinct in both cultures, as are the male voice profiles of sincere brands. Concerning exciting brands, the German male brand voice profile is the opposite of the American male profile. The findings of this cross-cultural comparison in vocal brand personality perception highlighted the necessity for organizations with an international presence to select or design a distinct brand voice depending on the brand gender and their customers' culture.

7.1. Managerial Implications

The importance of incorporating the cultural context into technological design is demonstrated by manufacturers of voice-based technologies, which are increasingly being developed in a culturally sensitive way to appeal to customers worldwide (Seaborn et al., 2024). For example, Amazon markets its smart speaker in 89 countries, offering voice services through Alexa in nine languages, including several regional accents and dialects (Amazon, 2024b). Moreover, Alexa adjusts its voice depending on the language set and can employ country-specific speech styles, such as an Australian-specific news delivery style (Gao, 2019). Google's virtual assistant, Google Home, can speak 16 languages and offers a range of voices tailored to a specific country or region. For instance, a voice with a higher pitch is available in Japan, reflecting the cultural preference for this particular voice style (Google, 2024; Starr, 2015). By offering regional dialects and accents, voice assistants are adjusted to a country's language on the vocal channel and on the verbal channel, which encodes the semantic content of a message (Apple et al., 1979; Kreiman & Sidtis, 2011). This adjustment uses country- and region-specific vocabulary, which may have different meanings depending on the context. One example is using the terms "fries" in American English and "chips" in British English to describe the same food.

The findings of this study suggest that managers should focus on the significant impact of culture on the perception of vocal brand personality dimensions, which is crucial for ensuring cultural sensitivity in voice AI technology. Even though international target audiences may exhibit similar communication patterns, due to the influence of learned norms within a culture, customers from different cultural groups attribute distinctive personalities to the same brand voice (Kreiman & Sidtis, 2001; Brown et al., 1975; Hogan & Bond, 2009). Thus, marketers engaging in voice marketing on an international scale are advised to consider the influence of culture in the selection of brand voice by referring to the derived brand voice profiles for Germany and America.

Although most comparisons between German and American brand voice profiles revealed distinct combinations of voice parameters inducing personality perceptions, notable similarities were observed in the case of sincere and exciting brands. In both cultures, a female voice shall be creaky and speak fluently. In Germany, the voice additionally needs to be breathy, while in the U.S., the voice needs to be rough. These are two voice quality parameters describing voices with either low laryngeal tension (breathiness) or high laryngeal tension (roughness; Barsties V Latoszek, Bodt de, et al., 2018; Dejonckere et al., 1993; Gobl & Chasaide, 2003).[11] Further, the voice profiles of exciting female brands in both cultures are strongly comparable regarding breathiness and creakiness in voice but slightly differ in timing parameters. In such cases, marketers may utilize the same voice in both languages and adjust single voice parameters for the respective country. Such voice modulations are typically the domain of professional sound engineers or voice designers. Additionally, phonetic software, plugins, and scripts can

[11] laryngeal tension = tightness in the muscles of the larynx (Clark et al. (2007)

be utilized, such as the PRAAT plugin Vocal Toolkit (Corretge, 2012-2024). This tool can introduce breathiness or raspiness (comparable to roughness or hoarseness) into a voice by, for instance, increasing the jitter.

In addition to adjusting voice parameters to the specific perceptions of a personality within a country, companies must tailor their voice to align with the characteristics of their target audience. This alignment ensures effective communication with customers via voice touchpoints. The findings of this study illustrate that especially the gender and age of the listener can significantly influence the perception of a brand's personality. Consequently, when international target audiences exhibit notable differences, it is advisable to adapt brand voices following the attributes of the individuals in the respective country.

The adaptation of a brand voice in accordance with the prevailing culture or target audience is a crucial aspect of international marketing and branding, as it promotes a consistent brand identity. Despite the necessity for regional adaptation strategies in specific branding aspects, for instance, in names or visual elements (Whitelock & Fastoso, 2007), consumers expect brands to possess a unified core concept on a global scale (Chernatony et al., 1995; Matthiesen & Phau, 2005). A consistent brand image has been demonstrated to provide numerous advantages, including a high level of brand awareness, the ability to leverage economies of scale in marketing communication, and an overall increase in brand equity (Bengtsson et al., 2010; Kapferer, 2011; L. E. Keller, 2013).

7.2. Theoretical Implications

This study investigated the vocal perception of brand personalities in an American English-speaking context and compared the findings with those from a German-speaking context. In doing so, it reacted to the two greatest limitations in the existing research on international vocal personality perception.

Firstly, this study makes a valuable contribution to the limited research on the vocal perception of brand personalities, as previous research focused on examining human personalities. As human personality traits do not necessarily apply to brands, considering appropriate brand personalities is essential. To ensure that the most relevant brand personality traits are investigated, the present study utilized items of three well-established scales for measuring a brand personality for the assessment of voices (Aaker, 1997; Geuens et al., 2009; Grohmann, 2009). The results of the confirmatory analyses indicated that none of the BPS can be reproduced. This finding suggested that certain predetermined brand personality traits cannot be conveyed solely through the voice. Therefore, an American BVP-Scale was developed to assess a brand's voice personality. The results of this study illustrated that in an American English-speaking context, the brand voice personality dimensions of sincerity, sensitivity, and excitement can be perceived through purely auditory communication.

Secondly, although little research addresses cultural differences in the perception of human personalities through voice, this study is the first to investigate the vocal perception of brand personalities and be situated within a cross-cultural context. The perception of personalities through vocal characteristics was found to be influenced by the cultural background of the listener (Charoenruk & Olson, 2018; Kreiman & Sidtis, 2011; H. O. Lee & Boster, 1992; van Bezooijen, 1995). Since social interaction norms and rules learned and followed in one culture are not necessarily the same as those in another, the same speaker may be assessed with different personality traits depending on the cultural context (Brown et al., 1975; Hogan & Bond, 2009). The comparison of the American and German BVP-Scales shows that similar brand voice personality dimensions are perceived in both cultures. In contrast, a comparison of the American and German BVP-Models reveals that the perception of the personality dimensions is induced through different linear combinations of voice parameters. These findings highlight the crucial role of culture in vocal personality perception and further extend the knowledge of how brand voices need to be evaluated when applied on an international scale.

Further, through its replication character, this study reacts to the general need for replicative research, especially in marketing science and consumer research (Easley et al., 2000; Urminsky & Dietvorst, 2024). Replicative research serves to confirm the reliability and generalizability of original research results, which is achieved by applying an identical or an alternative research design or source of data (e.g., sample, population; Dau et al., 2022; Nosek et al., 2022). Replication studies verify, consolidate, or extend previous research findings, which leads to a deeper understanding of knowledge and derived theories (Schmidt, 2009). As replication studies provide empirical evidence to support the generalizability of findings, there has been a growing emphasis on the importance of such studies across various academic disciplines in recent years (Dau et al., 2022; Easley et al., 2000; Kwon et al., 2017; Urminsky & Dietvorst, 2024). The replication character of this study is defined by its alignment with the German methodological and analytical procedures. This cross-cultural replication study, therefore, contributes to a more comprehensive understanding of the cultural influence on vocal brand personality perception. Furthermore, it contributes to the applicability of the utilized methods and models to facilitate the development of global strategies on vocal brand personality perception.

7.3. Directions for Future Research

This study aimed to examine how personality is perceived through voice and identify the influence of culture in this perceptual process. Therefore, a perceptual study was conducted in the United States to compare the results with the preceding study conducted in Germany. An American sample was chosen due to the prominence of voice technology in the United States (Triton Digital, 2022; Vixen Labs & Digitalscouting, 2022). Despite the similarities in communication patterns and styles of the German and American English cultures (both are low-context and linear-active cultures), the findings revealed that there are, in fact, substantial differences in the decisive voice parameters that induce perceptions of brand voice personality dimensions. This observation raises the question of how much brand voice profiles may differ between cultures exhibiting even greater values and communication differences. How is the same brand voice personality perceived in a low vs. high-context culture? Or in a linear-active vs. multi-active vs. reactive culture? Are there cultures similar in vocal brand personality perception? If so, what are the determinants of the similarities/differences? It is recommended that these questions be the focus of further replication studies conducted in other cultural contexts.

While this study focused on a cross-cultural comparison of vocal brand personality perception, it also identified potential avenues for future research that are more general in scope. In conjunction with previous research, the present study's findings guided the formulation of the following research directions.

Regional influences within the same culture can impact varying personality perceptions based on accents or dialects. For instance, individuals from Buffalo, New York, were perceived as more cordial, less potent, and less active than those with a New York City dialect (Markel et al., 1967). The different personality judgments observed can be attributed to the association of dialects with speech communities, reflecting regional origin and socioeconomic status (Krauss et al., 2002). Over time, stereotypes have developed that are recalled when hearing a dialect, which can also result in stigmatization and discrimination of individuals (Reinares-Lara et al., 2016). For instance, speakers associated with lower working-class backgrounds were judged to possess less favorable characteristics than middle-class speakers (Krauss et al., 2002; Smedley & Bayton, 1978). Because accents and dialects serve as vocal cues in transmitting the spokesperson's personality, future research is advised to investigate how accents and dialects affect the perception of brand personalities through voice (Reinares-Lara et al., 2016). Investigating the impact of accent and dialect in voice marketing can enhance the country of origin (COO) research. In branding, the communication of the COO (typically operationalized by the phrase "made in _") has been shown to evoke positive values and emotions of the respective origin, which shall positively influence product evaluations (Peterson & Jolibert, 1995; Puzakova et al., 2015; Verlegh & Steenkamp, 1999). The interaction between brand and consumer using voice-based devices is typically purely auditory, so the COO and a desired personality can be conveyed through brand voices with distinctive accents.

Recently, voice assistants like Alexa and Google Home were enabled to be empathic, which means that a conversational agent is capable of recognizing a consumer's emotional state and modifying their response based on that information (Asada, 2015; Mari et al., 2024; Rubio-Licht, 2023). For example, in case a user says, "I am stressed, are there any anxiety support groups in the area?" Google Home may respond with "I'm sorry to hear that" (Plauché & Berman, 2022; Rubio-Licht, 2023). Such AI empathy is based on speech recognition technology, in which the conversational agent processes the user's audio data into a transcript of the spoken word. This transcription is then screened for "one or more words" (Plauché & Berman, 2022, 002 background) that indicate an emotional need of the user. Thus, emotions are identified on a semantic level, i.e., through the verbal channel. The findings of previous research on human personality perception and the results of the present and preceding studies on the perception of brand personality could help promote research in detecting emotional states and personality traits through the vocal channel. To enable voice assistants to identify the personality of its users would be beneficial for brands as research has shown that consumers prefer to choose brands and prefer advertisement messages that match their personality (Bosnjak et al., 2007; Shumanov et al., 2022). The similarity-attraction theory can explain this behavior; humans respond more positively to people who are similar in personality, which was observed to hold between humans and robots and conversational agents (Andrist et al., 2015; Nass & Lee, 2001; Park et al., 2020; Reeves & Nass, 1996). Thus, further research on brand-consumer personality assimilation is needed to enhance the knowledge of improving the user experience and increasing shareholder value (Shumanov et al., 2022).

The final recommendation for future research on the vocal perception of brand personalities is to investigate the advertising and branding effectiveness of the BVP-Models. Brand voices developed following the derived voice profiles could be evaluated compared to the default voices of a voice assistant, such as Alexa, Google Home, or Siri. In this regard, potential impact variables that could be considered and measured include cognitive, affective, and conative ad and brand attitude (H. Lee & Cho, 2020; Martín-Santana et al., 2015), brand recall or purchase intent (Leung & Kee, 1999). Comparable studies were undertaken by investigating which spokesperson voices are preferred, enhance product recalls, or increase ad efficiency within the domains of telemarketing, direct selling, or radio and TV advertisement (Dahl, 2010; Ketrow, 1990; Martín-Santana et al., 2015; Rodero, 2017). Primarily, such research focused on the timing or frequency dimension of a spokesperson's voice, as voice parameters of these dimensions were the easiest to measure with early software for phonetic and speech analysis (Chattopadhyay et al., 2003; Labarbera & Maclachlan, 1979; Peterson et al., 1995; Poon et al., 2018; Sharf & Lehman, 1984; Skinner et al., 1999). Future research could benefit from building upon the existing findings on the efficacy of voices in traditional marketing communication by extending the analysis to voice marketing and voice-based AI domains. One possible approach would be to increase the number of voice parameters to be analyzed and apply multivariate analyses of parameters, as demonstrated in the present study.

Essay 3 References

Aaker, J. L. (1997). Dimensions of Brand Personality. *Journal of Marketing Research*, 34(3), 347–356. https://doi.org/10.1177/002224379703400304

Addington, D. W. (1968). The relationship of selected vocal characteristics to personality perception. *Speech Monographs*, 35(4), 492–503. https://doi.org/10.1080/03637756809375599

Amazon. (2024). *List of Alexa Interfaces and Supported Languages*. https://developer.amazon.com/en-US/docs/alexa/device-apis/list-of-interfaces.html

Anderson, R. C., Klofstad, C. A., Mayew, W. J., & Venkatachalam, M. (2014). Vocal fry may undermine the success of young women in the labor market. *PLOS ONE*, 9(5), e97506. https://doi.org/10.1371/journal.pone.0097506

Andrist, S., Mutlu, B., & Tapus, A. (2015). Look like me: matching robot personality via gaze to increase motivation. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 3603–3612). https://doi.org/10.1145/2702123.2702592

Apple, W., Streeter, L. A., & Krauss, R. M. (1979). Effects of pitch and speech rate on personal attributions. *Journal of Personality and Social Psychology*, 37(5), 715–727. https://doi.org/10.1037/0022-3514.37.5.715

Aronovitch, C. D. (1976). The voice of personality: Stereotyped judgments and their relation to voice quality and sex of speaker. *The Journal of Social Psychology*, 99(2), 207–220. https://doi.org/10.1080/00224545.1976.9924774

Asada, M. (2015). Development of artificial empathy. *Neuroscience Research*, 90, 41–50. https://doi.org/10.1016/j.neures.2014.12.002

Azoulay, A., & Kapferer, J.-N. (2003). Do brand personality scales really measure brand personality? *Journal of Brand Management*, 11(2), 143–155. https://doi.org/10.1057/palgrave.bm.2540162

Backhaus, K., Erichson, B., Gensler, S., Weiber, R., & Weiber, T. (2021). *Multivariate Analysemethoden*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-32425-4

Backhaus, K., Erichson, B., & Weiber, R. (2015). *Fortgeschrittene Multivariate Analysemethoden*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-46087-0

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16(1), 74–94. https://doi.org/10.1007/BF02723327

Barsties V Latoszek, B., Bodt de, M., Gerrits, E., & Maryn, Y. (2018). The Exploration of an Objective Model for Roughness With Several Acoustic Markers. *Journal of Voice: Official Journal of the Voice Foundation*, 32(2), 149–161. https://doi.org/10.1016/j.jvoice.2017.04.017

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Statistical Psychology*, 3(2), 77–85. https://doi.org/10.1111/j.2044-8317.1950.tb00285.x

Belin, P., Bestelmeyer, P. E. G., Latinus, M., & Watson, R. (2011). Understanding voice perception. *British Journal of Psychology*, 102(4), 711–725. https://doi.org/10.1111/j.2044-8295.2011.02041.x

Bengtsson, A., Bardhi, F., & Venkatraman, M. (2010). How global brands travel with consumers. *International Marketing Review*, 27(5), 519–540. https://doi.org/10.1108/02651331011076572

Boersma, P., & Weenink, D. (1992-2022). *Praat: doing phonetics by computer (Version 6.1.50)* [Computer software]. https://www.praat.org

Bosnjak, M., Bochmann, V., & Hufschmidt, T. (2007). Dimensions of brand personality attributions: A person-centric approach in the German cultural context. *Social Behavior and Personality: An International Journal*, 35(3), 303–316. https://doi.org/10.2224/sbp.2007.35.3.303

Bovens, M. (2007). Analysing and Assessing Accountability: A Conceptual Framework. *European Law Journal*, 13(4), 447–468. https://doi.org/10.1111/j.1468-0386.2007.00378.x

Brakus, J. J., Schmitt, B. H., & Zarantonello, L. (2009). Brand Experience: What is It? How is it Measured? Does it Affect Loyalty? *Journal of Marketing*, 73(3), 52–68. https://doi.org/10.1509/jmkg.73.3.052

Brown, B. L., Giles, H., & Thakerar, J. N. (1985). Speaker evaluations as a function of speech rate, accent and context. *Language & Communication*, 5(3), 207–220. https://doi.org/10.1016/0271-5309(85)90011-4

Brown, B. L., Strong, W. J., & Rencher, A. C. (1975). Acoustic determinants of perceptions of personality from speech. *Linguistics*, 13(166), 11–32. https://doi.org/10.1515/ling.1975.13.166.11

Caprara, G. V., Barbaranelli, C., & Guido, G. (2001). Brand personality: How to make the metaphor fit? *Journal of Economic Psychology*, 22(3), 377–395. https://doi.org/10.1016/s0167-4870(01)00039-3

Charoenruk, N., & Olson, K. (2018). Do Listeners Perceive Interviewers' Attributes from their Voices and Do Perceptions Differ by Question Type? *Field Methods*, 30(4), 312–328. https://doi.org/10.1177/1525822X18784500

Chattopadhyay, A., Dahl, D. W., Ritchie, R. J., & Shahin, K. N. (2003). Hearing Voices: The Impact of Announcer Speech Characteristics on Consumer Response to Broadcast Advertising. *Journal of Consumer Psychology*, 13(3), 198–204. https://doi.org/10.1207/S15327663JCP1303_02

Chen, A., Rietveld, T., & Gussenhoven, C. (2001). Language-specific effects of pitch range on the perception of universal intonational meaning. In *7th European Conference on Speech Communication and Technology (Eurospeech 2001)* (pp. 1403–1406). https://doi.org/10.21437/Eurospeech.2001-362

Chernatony, L. de, Halliburton, C., & Bernath, R. (1995). International branding: Demand– or supply-driven opportunity? *International Marketing Review*, 12(2), 9–21. https://doi.org/10.1108/02651339510089765

Clark, J., Yallop, C., & Fletcher, J. (2007). *An introduction to phonetics and phonology* (3rd ed.). Wiley-Blackwell.

Clopper, C. G., Levi, S. V., & Pisoni, D. B. (2006). Perceptual similarity of regional dialects of American English. *Journal of the Acoustical Society of America*, 119(1), 566–574. https://doi.org/10.1121/1.2141171

Corretge, R. (2012-2024). *Praat Vocal Toolkit.* [Praat Plugin]. https://www.praatvocal-toolkit.com

Coupé, C., Oh, Y. M., Dediu, D., & Pellegrino, F. (2019). Different languages, similar encoding efficiency: Comparable information rates across the human communicative niche. *Science Advances*, 5(9). https://doi.org/10.1126/sciadv.aaw2594

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. https://doi.org/10.1007/BF02310555

Dahl, D. W. (2010). Understanding the role of spokesperson voice in broadcast advertising. In A. Krishna (Ed.), *Sensory marketing: Research on the Sensuality of Products* (pp. 169–182). Routledge. https://doi.org/10.4324/9780203892060-21

Dau, L. A., Santangelo, G. D., & van Witteloostuijn, A. (2022). Replication studies in international business. *Journal of International Business Studies*, 53(2), 215–230. https://doi.org/10.1057/s41267-021-00471-w

Dejonckere, P. H., Obbens, C., Moor, G. M. de, & Wieneke, G. H. (1993). Perceptual evaluation of dysphonia: Reliability and relevance. *Folia Phoniatrica*, 45(2), 76–83. https://doi.org/10.1159/000266220

D'Errico, F., Signorello, R., Demolin, D., & Poggi, I. (2013). The Perception of Charisma from Voice: A Cross-Cultural Study. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction* (pp. 552–557). https://doi.org/10.1109/acii.2013.97

DeVellis, R. F. (2017). *Scale Development: Theory and Applications.* (4th ed.). SAGE Publications.

Dziuban, C. D., & Shirkey, E. C. (1974). When is a correlation matrix appropriate for factor analysis? Some decision rules. *Psychological Bulletin*, 81(6), 358–361. https://doi.org/10.1037/h0036316

Easley, R. W., Madden, C. S., & Dunn, M. G. (2000). Conducting Marketing Science. *Journal of Business Research*, 48(1), 83–92. https://doi.org/10.1016/S0148-2963(98)00079-4

eMarketer. Voice assistant users in the United States from 2022 to 2026 (in millions) [Graph]. In *Statista*. https://www.statista.com/statistics/1384575/voice-assistant-users-united-states/

Fenk-Oczlon, G., & Fenk, A. (2010). Measuring basic tempo across languages and some implications for speech rhythm. In *Interspeech 2010* (pp. 1537-1540). https://doi.org/10.21437/interspeech.2010-448

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39–50. https://doi.org/10.1177/002224378101800104

Gao, C. (2019). Use New Alexa Emotions and Speaking Styles to Create a More Natural and Intuitive Voice Experience. *Amazon*. https://developer.amazon.com/en-US/blogs/alexa/alexa-skills-kit/2019/11/new-alexa-emotions-and-speaking-styles

Geuens, M., Weijters, B., & Wulf, K. de (2009). A new measure of brand personality. *International Journal of Research in Marketing*, 26(2), 97–107. https://doi.org/10.1016/j.ijresmar.2008.12.002

Gobl, C., & Chasaide, A. N. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, 40(1-2), 189–212. https://doi.org/10.1016/S0167-6393(02)00082-1

Google. (2024). *Change the language of Google Assistant*. https://support.google.com/googlenest/answer/7550584?co=GENIE.Platform%3DAndroid&hl=en#zippy=%2Cgoogle-home%2Cgoogle-nest-hub%2Cspeakers

Grohmann, B. (2009). Gender Dimensions of Brand Personality. *Journal of Marketing Research*, 46(1), 105–119. https://doi.org/10.1509/jmkr.46.1.105

Ha, S., & Stoel, L. (2009). Consumer e-shopping acceptance: Antecedents in a technology acceptance model. *Journal of Business Research*, 62(5), 565–571. https://doi.org/10.1016/j.jbusres.2008.06.016

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2013). *Multivariate data analysis* (7th ed.). Pearson Education Limited.

Hair, J. F., Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, 109, 101–110. https://doi.org/10.1016/j.jbusres.2019.11.069

Hall, E. T. (1959). *The Silent Language*. Doubleday.

Hall, E. T. (1976). *Beyond Culture*. Doubleday.

Hall, E. T., & Hall, M. R. (2006). *Understanding Cultural Differences: Germans, French and Americans* (15th ed.). Intercultural Press.

Hess, T., Matt, C., Benlian, A., & Wiesböck, F. (2020). Options for Formulating a Digital Transformation Strategy. In R. Galliers, D. E. Leidner, & B. Simeonova (Eds.), *Strategic information management: Theory and practice* (5th edition, pp. 151–173). Routledge Taylor & Francis Group. https://doi.org/10.4324/9780429286797-7

Hildebrand, C., Efthymiou, F., Busquet, F., Hampton, W. H., Hoffman, D. L., & Novak, T. P. (2020). Voice analytics in business research: Conceptual foundations, acoustic feature extraction, and applications. *Journal of Business Research*, 121, 364–374. https://doi.org/10.1016/j.jbusres.2020.09.020

Hirst, D., & Di Cristo, A. (Eds.). (1998). *Intonation systems: A Survey of Twenty Languages*. Cambridge University Press.

Hofstede, G. (2011). Dimensionalizing Cultures: The Hofstede Model in Context. *Online Readings in Psychology and Culture,* 2(1). https://doi.org/10.9707/2307-0919.1014

Hogan, R., & Bond, M. H. (2009). Culture and personality. In P. Corr & G. Matthews (Eds.), *The Cambridge Handbook of Personality Psychology* (pp. 577–588). Cambridge: Cambridge University Press.

House, J. (2018). Communication Styles in English and German. In Y. Y. Kim (Ed.), *The Wiley Blackwell - ICA. The international encyclopedia of intercultural communication* (pp. 1–10). Wiley-Blackwell. https://doi.org/10.1002/9781118783665.ieicc0240

House, J., & Kasper, G. (1981). Politeness Markers in English and German. In F. Coulmas (Ed.), *Conversational Routine* (pp. 157–186). https://doi.org/10.1515/9783110809145.157

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multi-disciplinary Journal*, 6(1), 1–55. https://doi.org/10.1080/10705519909540118

Johnstone, t., & Scherer, K. R. (2000). Vocal communication of emotion. In Lewis M. & Haviland J. (Eds.), *Handbook of Emotion* (2nd ed.). New York: Guilford.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31–36. https://doi.org/10.1007/BF02291575

Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, 34(1), 111–117. https://doi.org/10.1177/001316447403400115

Kapferer, J.-N. (2011). *The new strategic brand management: Creating and sustaining brand equity long term* (4. ed., reprinted.). Kogan Page.

Keller, E. (2005). The Analysis of Voice Quality in Speech Processing. In G. Chollet, A. Esposito, M. Faundez-Zanuy, & M. Marinaro (Eds.), *Lecture Notes in Computer Science: Vol. 3445, Nonlinear Speech Modeling and Applications: Advanced Lectures and Revised Selected Papers* (pp. 54–73). Springer-Verlag GmbH. https://doi.org/10.1007/11520153_4

Keller, L. E. (2013). *Strategic Brand Management: Building, Measuring, and Managing Brand Equity*. UK: Pearson Education Limited.

Ketrow, S. M. (1990). Attributes of a telemarketer's voice and persuasiveness.: A review and synthesis of the literature. *Journal of Direct Marketing*, 4(3), 7–21. https://doi.org/10.1002/dir.4000040304

Kiesling, S., Dilley, L., & Raymond, W. D. (2006). The Variation in Conversation (ViC) Project: Creation of the Buckeye Corpus of Conversational Speech: Buckeye Corpus Manual. *Language Variation and Change*, 55-97.

Klasmeyer, G., & Sendlmeier, W. F. (1997). The classification of different phonation types in emotional and neutral speech. *International Journal of Speech, Language and the Law*, 4(1), 104–124. https://doi.org/10.1558/ijsll.v4i1.104

Korac-Kakabadse, N., Kouzmin, A., Korac-Kakabadse, A., & Savery, L. (2001). Low- and high-context communication patterns: Towards mapping cross-cultural encounters. *Cross Cultural Management: An International Journal*, 8(2), 3–24. https://doi.org/10.1108/13527600110797218

Koutsombogera, M., Sarthy, P., & Vogel, C. (2020). Acoustic Features in Dialogue Dominate Accurate Personality Trait Classification. In *2020 IEEE International Conference on Human-Machine Systems (ICHMS)* (pp. 1–3). https://doi.org/10.1109/ichms49158.2020.9209445

Krahé, B., Uhlmann, A., & Herzberg, M. (2021). The Voice Gives It Away: Male and Female Pitch as a Cue for Gender Stereotyping. *Social Psychology*, 52(2), 101–113. https://doi.org/10.1027/1864-9335/a000441

Krauss, R. M., Freyberg, R., & Morsella, E. (2002). Inferring speakers' physical attributes from their voices. *Journal of Experimental Social Psychology*, 38(6), 618–625. https://doi.org/10.1016/S0022-1031(02)00510-3

Kreiman, J., & Sidtis, D. (2011). Perception of Emotion and Personality from Voice. In *Foundations of Voice Studies* (pp. 302–360). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444395068.ch9

Kwon, E. S., Shan, Y., Lee, J. S., & Reid, L. N. (2017). Inter-study and intra-study replications in leading marketing journals: A longitudinal analysis. *European Journal of Marketing*, 51(1), 257–278. https://doi.org/10.1108/EJM-07-2015-0450

Labarbera, P., & Maclachlan, J. (1979). Time-Compressed Speech in Radio Advertising. *Journal of Marketing*, 43(1), 30–36. https://doi.org/10.1177/002224297904300103

Labov, W. (1991). The Three Dialects of English. In Penelope Eckert (Ed.), *New Ways of Analyzing Sound Change* (Vol. 5, pp. 1–44). Academic Press. http://ereserve.library.utah.edu/annual/ling/6040/dipaolo/ling60403dialects.pdf

Lee, H., & Cho, C.-H. (2020). Uses and gratifications of smart speakers: Modelling the effectiveness of smart speaker advertising. *International Journal of Advertising*, 39(7), 1150–1171. https://doi.org/10.1080/02650487.2020.1765657

Lee, H. O., & Boster, F. J. (1992). Collectivism-Individualism in Perceptions of Speech Rate. *Journal of Cross-Cultural Psychology*, 23(3), 377–388. https://doi.org/10.1177/0022022192233008

Leung, L., & Kee, O. K. (1999). The effects of male celebrity voice-over and gender on product brand name recall, comprehension, and purchase intention. *New Jersey Journal of Communication*, 7(1), 81–92. https://doi.org/10.1080/15456879909367356

Lewis, R. D. (2006). *When cultures collide: Leading across cultures* (3rd ed.). Nicholas Brealey International.

Lorenzo-Seva, U. (2021). Solomon: A method for splitting a sample into equivalent subsamples in factor analysis. *Behavior Research Methods*, 54(6), 2665–2677. https://doi.org/10.3758/s13428-021-01750-y

Mairesse, F., & Walker, M. (2006). Words Mark the Nerds: Computational Models of Personality Recognition through Language. In *Proceedings of the Annual Meeting 2006* (Vol. 28, pp. 543–548).

Mari, A., Mandelli, A., & Algesheimer, R. (2024). Empathic voice assistants: Enhancing consumer responses in voice commerce. *Journal of Business Research*, 175, 114566. https://doi.org/10.1016/j.jbusres.2024.114566

Markel, N. N., Eisler, R. M., & Reese, H. W. (1967). Judging personality from dialect. *Journal of Verbal Learning and Verbal Behavior*, 6(1), 33–35. https://doi.org/10.1016/S0022-5371(67)80044-6

Martín-Santana, J. D., Muela-Molina, C., Reinares-Lara, E., & Rodríguez-Guerra, M. (2015). Effectiveness of radio spokesperson's gender, vocal pitch and accent and the use of music in radio advertising. *BRQ Business Research Quarterly*, 18(3), 143–160. https://doi.org/10.1016/j.brq.2014.06.001

Matthiesen, I., & Phau, I. (2005). The ‘HUGO BOSS’ connection: Achieving global brand consistency across countries. *Journal of Brand Management*, 12(5), 325–338. https://doi.org/10.1057/palgrave.bm.2540229

Mayer, J. (2019-2021). *Praat-Stimmprofil: v2.3.0* [Praat Plugin].

McAleer, P., & Belin, P. (2019). The Perception of Personality Traits from Voices. In S. Frühholz & P. Belin (Eds.), *The Oxford Handbook of Voice Perception* (1st ed., pp. 585–606). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198743187.013.26

McAleer, P., Todorov, A., & Belin, P. (2014). How Do You Say 'Hello'? Personality Impressions from Brief Novel Voices. *PLOS ONE*, 9(3), 1–9. https://doi.org/10.1371/journal.pone.0090779.t001

McCrae, R. R., Costa, P. T., & Martin, T. A. (2005). The NEO-PI-3: A more readable revised NEO Personality Inventory. *Journal of Personality Assessment*, 84(3), 261–270. https://doi.org/10.1207/s15327752jpa8403_05

Mennen, I., Schaeffler, F., & Docherty, G. (2012). Cross-language differences in fundamental frequency range: A comparison of English and German. *The Journal of the Acoustical Society of America*, 131(3), 2249–2260. https://doi.org/10.1121/1.3681950

Mertens, P. (2022). The Prosogram Model for Pitch Stylization and Its Applications in Intonation Transcription. In J. A. Barnes & S. Shattuck-Hufnagel (Eds.), *Prosodic Theory and Practice* (pp. 259–286). MA: MIT Press. https://doi.org/10.7551/mit-press/10413.003.0010

Michalsky, J., Niebuhr, O., & Penke, L. (2020). Do charismatic people produce charismatic speech? On the relationship between the Big Five personality traits and prosodic features of speaker charisma in female speakers. In *10th International Conference on Speech* (pp. 700–704). https://doi.org/10.21437/SpeechProsody.2020-143

Nam, J., Ekinci, Y., & Whyatt, G. (2011). Brand equity, brand loyalty and consumer satisfaction. *Annals of Tourism Research*, 38(3), 1009–1030. https://doi.org/10.1016/j.annals.2011.01.015

Nass, C., & Lee, K. M. (2001). Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology: Applied*, 7(3), 171–181. https://doi.org/10.1037/1076-898x.7.3.171

Nishimura, S., Nevgi, A., & Tella, S. (2008). Communication style and cultural features in high/low context communication cultures: A case study of Finland, Japan and IndiaNANo ranking found for 'The Cultural Context in Business Communication'. In Kallioniemi, A. T. (Ed.), *Uudistuva ja kehittyvä ainedidaktiikka. Ainedidaktinen symposiumi* (pp. 783-796).

Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fid-ler, F., Hilgard, J., Kline Struhl, M., Nuijten, M. B., Rohrer, J. M., Romero, F., Scheel, A. M., Scherer, L. D., Schönbrodt, F. D., & Vazire, S. (2022). Replicability, Robustness, and Reproducibility in Psychological Science. *Annual Review of Psychology*, 73(Volume 73, 2022), 719–748. https://doi.org/10.1146/annurev-psych-020821-114157

Oksenberg, L., Coleman, L., & Cannell, C. F. (1986). Interviewers' Voices and Refusal Rates in Telephone Surveys. *Public Opinion Quarterly*, 50(1), 97. https://doi.org/10.1086/268962

Pang, E. (2021). Exploring Cultural Awareness When Designing Voice UX: Are voice assistants culturally sensitive? *Medium*. https://medium.com/55-minutes/the-importance-of-cultural-awareness-in-design-14a5c0ae2ad8

Park, J [Jangwoon], Lee, S., Brotherton, K., Um, D., & Park, J [Jaehyun] (2020). Identification of Speech Characteristics to Distinguish Human Personality of Introversive and Extroversive Male Groups. *International Journal of Environmental Research and Public Health,* 17(6). https://doi.org/10.3390/ijerph17062125

Paul Roberge. (2020). Contact and the History of Germanic Languages. In *The Handbook of Language Contact* (pp. 323–343). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781119485094.ch16

Pellegrino, F., Coupé, C., & Marsico, E. (2011). Across-Language Perspective on Speech Information Rate. *Language*, 87(3), 539–558. https://doi.org/10.1353/lan.2011.0057

Peng, Y., Zebrowitz, L. A., & Lee, H. K. (1993). The Impact of Cultural Background and Cross-Cultural Experience on Impressions of American and Korean Male Speakers. *Journal of Cross-Cultural Psychology*, 24(2), 203–220. https://doi.org/10.1177/0022022193242005

Peterson, R. A., Cannito, M. P., & Brown, S. P. (1995). An Exploratory Investigation of Voice Characteristics and Selling Effectiveness. *Journal of Personal Selling & Sales Management*, 15(1), 1–15. https://doi.org/10.1080/08853134.1995.10754008

Peterson, R. A., & Jolibert, A. J. P. (1995). A Meta-Analysis of Country-of-Origin Effects. *Journal of International Business Studies*, 26(4), 883–900. https://doi.org/10.1057/palgrave.jibs.8490824

Pitt, M. A., Dilley, L., Johnson, K., Kiesling, S., Raymond, W., Hume, E., & Fosler-Lussier, E. (2007). *Buckeye Corpus of Conversational Speech* (2nd release). Columbus, OH: Department of Psychology, Ohio State University (Distributor) [www.buckeyecorpus.osu.edu]

Plauché, M., & Berman, K. B. (2022). *Emotionally Intelligent Responses To Information Seeking Questions* (US 20230298580 A1). United States Patent and Trademark Office.

Poon, M. M., Chan, K. M., & Yiu, E. M. (2018). The Relationship Between Speech Rate, Voice Quality and Listeners' Purchase Intentions. In *9th International Conference on Speech Prosody 2018* (pp. 468-472). https://doi.org/10.21437/speechprosody.2018-95

Puzakova, M., Kwak, H., & Bell, M. (2015). Beyond Seeing McDonald's Fiesta Menu: The Role of Accent in Brand Sincerity of Ethnic Products and Brands. *Journal of Advertising*, 44(3), 219–231. https://doi.org/10.1080/00913367.2014.957367

Ray, G. B. (1986). Vocally cued personality prototypes: An implicit personality theory approach. *Communication Monographs*, 53(3), 266–276. https://doi.org/10.1080/03637758609376141

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people*. Cambridge University Press.

Reinares-Lara, E., Martín-Santana, J. D., & Muela-Molina, C. (2016). The Effects of Accent, Differentiation, and Stigmatization on Spokesperson Credibility in Radio Advertising. *Journal of Global Marketing*, 29(1), 15–28. https://doi.org/10.1080/08911762.2015.1119919

Revelle, W. (2024). *psych: Procedures for Psychological, Psychometric, and Personality Research (Version 2.4.3)* [Computer software]. https://CRAN.R-project.org/package=psych

Riding, D., Lonsdale, D., & Brown, B. (2006). The Effects of Average Fundamental Frequency and Variance of Fundamental Frequency on Male Vocal Attractiveness to Women. *Journal of Nonverbal Behavior*, 30(2), 55–61. https://doi.org/10.1007/s10919-006-0005-3

Rodero, E. (2013). The Perception of a Broadcasting Voice. *US-China Education Review*, 3(4), 225–230.

Rodero, E. (2017). Effectiveness, attention, and recall of human and artificial voices in an advertising story. Prosody influence and functions of voices. *Computers in Human Behavior*, 77, 336–346. https://doi.org/10.1016/j.chb.2017.08.044

Rosseel, Y. (2012). Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software,* 48(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Rubio-Licht, N. (2023). Google is Adding Empathy to it's Smart Speakers. *The Daily Upside*. https://www.thedailyupside.com/technology/big-tech/google-is-adding-empathy-to-its-smart-speakers/

Sarstedt, M., & Mooi, E. (2019). Regression Analysis. In *A Concise Guide to Market Research* (pp. 209–256). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-56707-4_7

Sarstedt, M., Ringle, C. M., & Hair, J. F. (2020). Partial Least Squares Structural Equation Modeling. In C. Homburg, M. Klarmann, & A. Vomberg (Eds.), *Springer eBook Collection Business and Management. Handbook of market research* (pp. 1–47). Springer. https://doi.org/10.1007/978-3-319-05542-8_15-2

Scherer, K. R. (1972). Judging personality from voice: A cross-cultural approach to an old issue in interpersonal perception. *Journal of Personality*, 40(2), 191–210. https://doi.org/10.1111/j.1467-6494.1972.tb00998.x

Scherer, K. R. (1974). Voice quality analysis of American and German speakers. *Journal of Psycholinguistic Research*, 3(3), 281–298. https://doi.org/10.1007/BF01069244

Scherer, K. R. (1979). Personality markers in speech. In Scherer, K. R., & Giles, H. (Ed.), *Social markers in speech* (Vol. 6, pp. 147–209). Cambridge University Press.

Schmidt, S. (2009). Shall we Really do it Again? The Powerful Concept of Replication is Neglected in the Social Sciences. *Review of General Psychology*, 13(2), 90–100. https://doi.org/10.1037/a0015108

Schwartz, E. H. (2019). Siri Can Tie and Untie Nike's New Sneakers. *voicebot.ai*. https://voicebot.ai/2019/09/04/siri-can-tie-and-untie-nikes-new-sneakers/

Schwartz, E. H. (2020). Australians Can Ask Alexa for a Personalized Bottle of Coke. *voicebot.ai*. https://voicebot.ai/2020/01/30/australians-can-ask-alexa-for-a-personalized-bottle-of-coke/

Seaborn, K., Miyake, N. P., Pennefather, P., & Otake-Matsuura, M. (2022). Voice in Human–Agent Interaction. *ACM Computing Surveys*, 54(4), 1–43. https://doi.org/10.1145/3386867

Seaborn, K., Sawa, Y., & Watanabe, M. (2024). Coimagining the Future of Voice Assistants with Cultural Sensitivity. *Human Behavior and Emerging Technologies*, 2024(1), 1–21. https://doi.org/10.1155/2024/3238737

Sharf, D. J., & Lehman, M. E. (1984). Relationship between the speech characteristics and effectiveness of telephone interviewers. *Journal of Phonetics*, 12(3), 219–228. https://doi.org/10.1016/S0095-4470(19)30878-2

Shumanov, M., Cooper, H., & Ewing, M. (2022). Using AI predicted personality to enhance advertising effectiveness. *European Journal of Marketing*, 56(6), 1590–1609. https://doi.org/10.1108/EJM-12-2019-0941

Skinner, C. H., Robinson, D. H., Robinson, S. L., Sterling, H. E., & Goodman, M. A. (1999). Effects of Advertisement Speech Rates on Feature Recognition, and Product and Speaker Ratings. *International Journal of Listening*, 13(1), 97–110. https://doi.org/10.1080/10904018.1999.10499029

Smedley, J. W., & Bayton, J. A. (1978). Evaluative race-class stereotypes by race and perceived class of subjects. *Journal of Personality and Social Psychology*, 36(5), 530–535. https://doi.org/10.1037//0022-3514.36.5.530

Starr, R. L. (2015). Sweet voice: The role of voice quality in a Japanese feminine style. *Language in Society*, 44(1), 1–34. https://doi.org/10.1017/S0047404514000724

Steenkamp, J.-B. E. M., & Baumgartner, H. (1998). Assessing Measurement Invariance in Cross-National Consumer Research. *Journal of Consumer Research*, 25(1), 78–107. https://doi.org/10.1086/209528

Tian, K. T., Bearden, W. O., & Hunter, G. L. (2001). Consumers' Need for Uniqueness: Scale Development and Validation. *Journal of Consumer Research*, 28(1), 50–66. https://doi.org/10.1086/321947

Triton Digital. (2022). Share of individuals who own any smart speaker brand in selected countries in 2020 and 2022, by country [Graph]. In *Statista*. https://www.statista.com/statistics/1130208/smart-speaker-ownership-worldwide-by-country/c

Trouvain, J., Weiss, B., & Barkat-Defradas, M. (2021). Voice Attractiveness: Concepts, Methods, and Data. In *Voice attractiveness: Studies on sexy, likable, and charismatic speakers* (pp. 3–16). Springer, Singapore. https://doi.org/10.1007/978-981-15-6627-1_1

Urminsky, O., & Dietvorst, B. J. (2024). Taking the Full Measure: Integrating Replication into Research Practice to Assess Generalizability. *Journal of Consumer Research*, 51(1), 157–168. https://doi.org/10.1093/jcr/ucae007

van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, 38(3), 253–265. https://doi.org/10.1177/002383099503800303

Verlegh, P. W., & Steenkamp, J.-B. E. (1999). A review and meta-analysis of country-of-origin research. *Journal of Economic Psychology*, 20(5), 521–546. https://doi.org/10.1016/S0167-4870(99)00023-9

Vixen Labs, & Digitalscouting. (2022). *Voice Consumer Index 2022*. https://vixenlabs.co/research/voice-consumer-index-2022

VML. (2023). *A built-in Alexa skill that rewards viewers every time a Coke pops on screen.* https://www.vml.com/work/i-see-coke

Volkswagen Group. (2024). *ChatGPT is now available in many Volkswagen models*. https://www.volkswagen-group.com/en/articles/chatgpt-is-now-available-in-many-volkswagen-models-18464

Waaramaa, T., Lukkarila, P., Järvinen, K., Geneid, A., & Laukkanen, A.-M. (2021). Impressions of Personality from Intentional Voice Quality in Arabic-Speaking and Native Finnish-Speaking Listeners. *Journal of Voice: Official Journal of the Voice Foundation*, 35(2), 326.e21-326.e28. https://doi.org/10.1016/j.jvoice.2019.09.003

Wang, L., Huang, N., Hong, Y., Liu, L., Guo, X., & Chen, G. (2023). Voice-based AI in call center customer service: A natural field experiment. *Production and Operations Management*, 32(4), 1002–1018. https://doi.org/10.1111/poms.13953

Weiber, R., & Mühlhaus, D. (2014). *Strukturgleichungsmodellierung*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-35012-2

Wenden, A. (2020). Alexa Speaking Styles and Emotions Now Available in Additional Languages. *Amazon*. https://developer.amazon.com/en-US/blogs/alexa/alexa-skills-kit/2020/11/alexa-speaking-styles-emotions-now-available-additional-languages

Westlund, A. H., Källström, M., & Parmler, J. (2008). SEM-based customer satisfaction measurement: On multicollinearity and robust PLS estimation. *Total Quality Management & Business Excellence,* 19(7-8), 855–869. https://doi.org/10.1080/14783360802159527

Whitelock, J., & Fastoso, F. (2007). Understanding international branding: Defining the domain and reviewing the literature. *International Marketing Review*, 24(3), 252–270. https://doi.org/10.1108/02651330710755285

Winn, M. B. (2020). Manipulation of voice onset time in speech stimuli: A tutorial and flexible Praat script. *The Journal of the Acoustical Society of America*, 147(2), 852. https://doi.org/10.1121/10.0000692

Yong, A. G., & Pearce, S. (2013). A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutorials in Quantitative Methods for Psychology*, 9(2), 79–94. https://doi.org/10.20982/tqmp.09.2.p079

Zäske, R., Skuk, V. G., Golle, J., & Schweinberger, S. R. (2020). The Jena Speaker Set (JESS)-A database of voice stimuli from unfamiliar young and old adult speakers. *Behavior Research Methods*, 52(3), 990–1007. https://doi.org/10.3758/s13428-019-01296-0

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99(3), 432–442. https://doi.org/10.1037//0033-2909.99.3.432

Essay 3 Appendix

Appendix A. Categorization & Definition of Voice Parameters

Voice Dimension	Listener's Perception	Acoustical Measure (Metric)	Definition
Timing	Fluency	Speaking Rate (syl/s)	Number of words or syllables spoken per unit of time, including voice breaks (Peterson et al., 1995; Tusing & Dillard, 2000).
		Average Silent Pause Duration (s) Silent Pause Frequency per Minute (n/m)	Silent pauses are unfilled pauses that may be accompanied by an inhalation, exhalation, or swallowing (Conrad et al., 2008; Koutsoumpis & Vries, 2022; Rodero, 2012).
	Velocity of Speech	Articulation Rate (syl/s)	Number of words or syllables spoken per unit of time, excluding voice breaks (Street & Brady, 1982).
Amplitude	Loudness Variability (Intonation)	Intensity Variability (SD; dB)	Sound intensity is measured in decibels (dB) and represents the amplitude of a sound wave per unit area (Hildebrand et al., 2020). The standard deviation of individual intensity levels indicates intonation in speech (Scherer, 1974).
Frequency	Pitch	Fundamental Frequency Mean (f0; Hz)	Frequency refers to a sound wave's frequency, measured in Hertz (Hz). The fundamental frequency (f0) describes the number of vibrations per second that the vocal folds make to produce a vocalization (Hildebrand et al., 2020; Koutsoumpis & Vries, 2022). SD f0 and range f0 refer to the rise and fall of f0 over an utterance and are perceived as pitch variability, which serves as indicators of intonation.
	Pitch Variability (Intonation)	Fundamental Frequency Standard Deviation (SD f0; st) Fundamental Frequency Range (f0-max – f0-min; Hz)	
Voice Quality	Creakiness/ Breathiness	h1-h2 (dB)	Amplitude difference between the 1st and 2nd harmonics, i.e., the relationship between open and closed phases of the glottis (Barsties V Latoszek, Maryn, et al., 2018).
		Spectral Slope*	The rate of amplitude decreases between two increasing frequencies in a spectrum** (McAleer et al., 2014).
		Spectral Tilt*	The overall degree to which intensity drops off as frequency increases, i.e., the slope of the trend line through the spectrum (M. Gordon & Ladefoged, 2001).
	Roughness/ Hoarseness	Harmonics-to-Noise Ratio (HNR; dB)	Relationship between the periodic (harmonics) and aperiodic (noise) components of a speech (McAleer et al., 2014).
		Jitter (%)	Short-term perturbation of fundamental frequency (f0; Clark et al., 2007).
		Shimmer (%)	Short-term perturbation of amplitude (Clark et al., 2007).
	Breathiness	Smoothed Cepstral Peak Prominence (CPPS; dB)	CPPS quantifies the prominence of the harmonic structure (periodic vocal fold vibrations) over noise in voice (Baker et al., 2022; Fraile & Godino-Llorente, 2014).

Notes: * Spectral slope and tilt are assigned to the voice quality dimension, even though these voice parameters are also indicators of perceived loudness. ** A spectrum represents the distribution of intensity based on frequency for a specific sound source. **Abbreviations:** syl = syllable; s = second; m = minute; n = number; SD = standard deviation; Hz = Hertz; st = semitones; dB = Decibel.

Appendix B. Voice Recordings & Speaker's Characteristics from the Buckeye Corpus

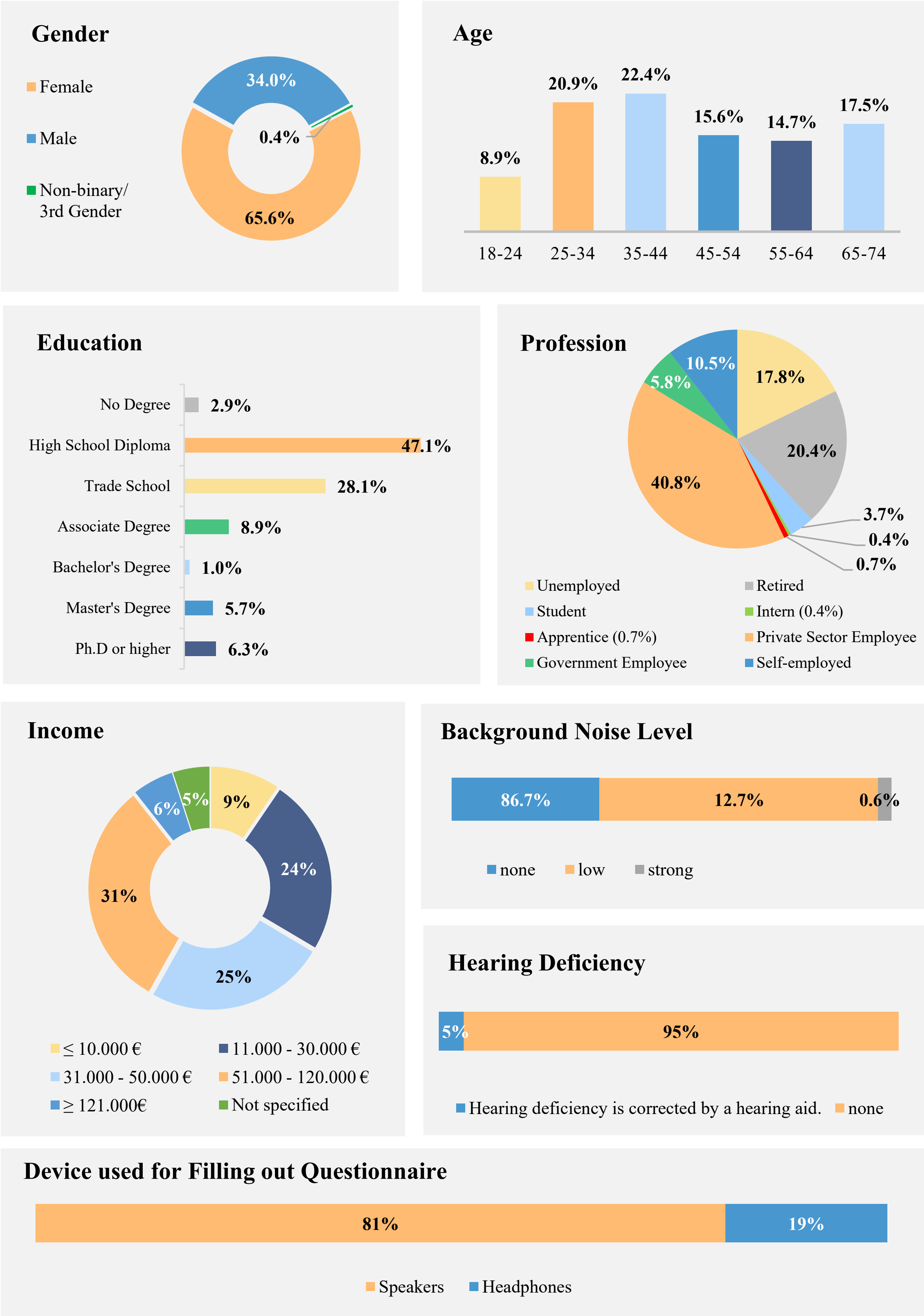
Speaker ID	Selected Voice Recording	Length of Voice Sample (seconds)	Speaker's Gender	Speaker's Age
S01	s0103a	17.67	female	under 30
S02	s0206a	17.19	female	over 40
S03	s0301b	15.05	male	over 40
S04	s0401a	13.68	female	under 30
S05	s0501a	17.43	female	over 40
S06	s0601a	15.06	male	under 30
S07	s0701a	20.37	female	over 40
S08	s0801a	17.35	female	under 30
S09	s0901a	17.75	female	under 30
S10	s1001a	19.39	male	over 40
S11	s1102b	17.91	male	under 30
S12	s1201a	15.38	female	under 30
S13	s1301a	17.76	male	under 30
S14	s1401a	16.16	female	over 40
S15	s1502b	19.48	male	under 30
S16	s1604a	17.50	female	over 40
S17	s1701a	17.02	female	over 40
S18	s1801a	17.65	female	over 40
S19	s1901a	18.76	male	over 40
S20	s2001b	16.56	female	over 40
S21	s2101a	11.97	female	under 30
S22	s2203b	19.39	male	over 40
S23	s2301a	15.61	male	over 40
S24	s2401b	16.78	male	over 40
S25	s2501a	20.07	female	over 40
S26	s2603a	15.18	female	under 30
S27	s2704a	18.88	female	over 40
S28	s2801a	20.63	male	under 30
S29	s2901a	15.42	male	over 40
S30	s3003a	19.37	male	under 30
S31	s3101b	12.13	female	under 30
S32	s3201a	11.41	male	under 30
S33	s3301b	17.41	male	under 30
S34	s3401a	20.47	male	under 30
S35	s3501a	18.59	male	over 40
S36	s3601a	13.01	male	over 40
S37	s3701a	18.71	female	under 30
S38	s3801b	20.16	male	over 40
S39	s3902a	16.80	female	under 30
S40	s4001a	14.99	male	under 30

Appendix C. Access to Vocal Stimulus Set of this Study

The 40 voices used as this study's vocal stimuli were obtained from the Buckeye corpus of conversational speech (Pitt et al., 2007). The interviews conducted for corpus creation lasted from 30 to 60 minutes. For this study, samples were cut out of each interview, which contained neutral content without emotions or controversial topics and had no interruptions through the interviewee, background noises, or linguistic breaks.

The vocal stimulus set of this study is provided in the electronic appendix of this dissertation – in the folder named "Essay_3_Voice_Samples_Buckeye". The folder contains 40 audio files, each named according to the declared **voice sample number for this study** and the sample's *original file name from the Buckeye corpus*, e.g., "**voicesample1**_*s0103a*".

Appendix D. Sample Characteristics: Socio-Demographic & Survey-Related Data



Notes: N = 979 participants.

Appendix E. Confirmatory Factor Analysis Aaker's (1997) Brand Personality Scale

Construct/ Item	Standardized Factor Loading (|≥.5|)	CR (>.7)	AVE (>.5)
Competence			
confident	.73***		
corporate	.54***		
hard working	.71***		
intelligent	.75***		
leader	.76***	.90	.50
reliable	.76***		
secure	.75***		
successful	.80***		
technical	.47***		
Excitement			
contemporary	.62***		
cool	.75***		
daring	.62***		
exciting	.74***		
imaginative	.72***		
independent	.65***	.89	.43
spirited	.70***		
trendy	.69***		
unique	.66***		
up-to-date	.66***		
young	.34***		
Ruggedness			
rugged	.75***		
masculine	.38***		
outdoorsy	.61***	.71	.34
tough	.69***		
western	.40***		
Sincerity			
cheerful	.63***		
down-to-earth	.68***		
family-oriented	.63***		
friendly	.74***		
honest	.71***		
original	.60***	.89	.42
real	.65***		
sentimental	.59***		
sincere	.74***		
small-town	.30***		
wholesome	.75***		
Sophistication			
charming	.75***		
feminine	.29***		
glamorous	.65***	.77	.64
good-looking	.65***		
smooth	.66***		
upper class	.58***		

Notes: Method: Maximum Likelihood Estimation with Robust Standard Errors and Satorra-Bentler (S-B) correction. Values below the specified thresholds are highlighted in red. Fornell-Lacker criterion not met (AVE > squared correlation of the latent construct with the discriminant construct). N = 1,958 individual voice ratings; *** p < .001.

Global Model Fit Indices: S-Bχ^2 (809) = 9,380.122 (p < .001); χ^2/df = 11.59; Standardized Root Mean Square Residual (SRMR) = .08; Root Mean Square Error of Approximation (RMSEA) = .07; Comparative Fit Index (CFI) = .80; Tucker-Lewis Index (TLI) = .78; Goodness-of-Fit Index (GFI) = .78; Adjusted Goodness of Fit Index (AGFI) = .75; Normed Fit Index (NFI) = .78.

Abbreviations: CR = Composite Reliability; AVE = Average Variance Extracted.

Appendix F. Confirmatory Factor Analysis Geuens et al.'s (2009) Brand Personality Scale

Construct/ Item	Standardized Factor Loading (|≥.5|)	CR (>.7)	AVE (>.5)
Activity			
dynamic	.75***	.76	.52
innovative	.73***		
active	.68***		
Simplicity			
simple	.71***	.52	.36
ordinary	.47***		
Responsibility			
down-to-earth	.60***	.75	.50
responsible	.77***		
stable	.75***		
Emotionality			
romantic	.56***	.53	.36
sentimental	.64***		
Aggressiveness			
bold	.92***	.61	.48
aggressive	.34***		

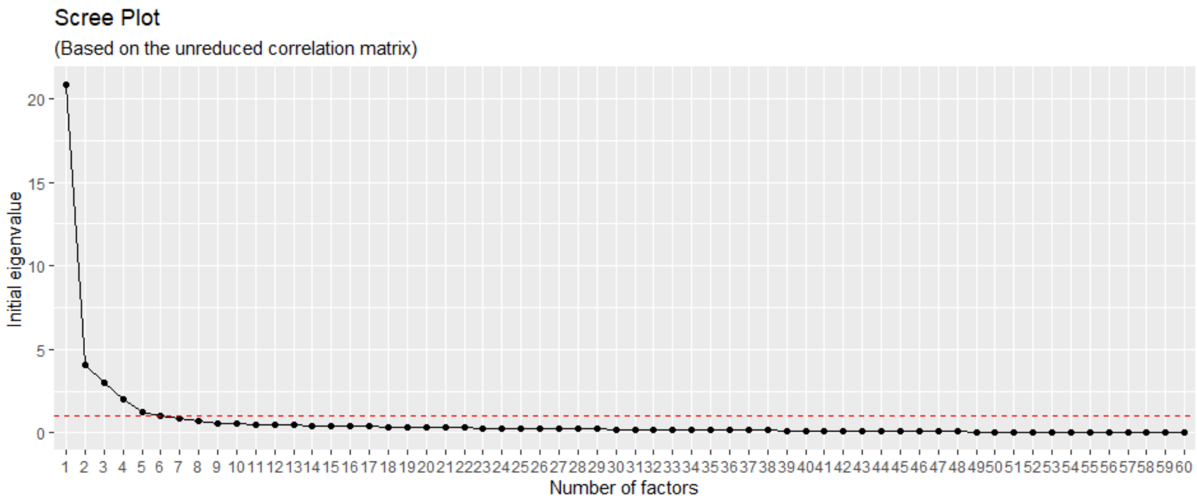
Notes: Method: Maximum Likelihood Estimation with Robust Standard Errors and Satorra-Bentler (S-B) correction. Values below the specified thresholds are highlighted in red. Fornell-Lacker criterion not met (AVE > squared correlation of the latent construct with the discriminant construct). N = 1,958 individual voice ratings; *** p < .001.
Global Model Fit Indices: S-Bχ^2 (44) = 381.511 (p < .001); χ^2/df = 8.67; Standardized Root Mean Square Residual (SRMR) = .04; Root Mean Square Error of Approximation (RMSEA) = .06; Comparative Fit Index (CFI) = .95; Tucker-Lewis Index (TLI) = .92; Goodness-of-Fit Index (GFI) = .97; Adjusted Goodness of Fit Index (AGFI) = .94; Normed Fit Index (NFI) = .94.
Abbreviations: CR = Composite Reliability; AVE = Average Variance Extracted.

Appendix G. Confirmatory Factor Analysis Grohmann's (2009) Brand Personality Scale

Construct/ Item	Standardized Factor Loading (|≥.5|)	CR (>.7)	AVE (>.5)
Feminity			
expresses tender feelings	.78***	.86	.50
fragile	.45***		
graceful	.68***		
sensitive	.73***		
sweet	.77***		
tender	.77***		
Masculinity			
adventurous	.67***	.78	.38
aggressive	.38***		
brave	.73***		
daring	.76***		
dominant	.61***		
rugged	.44***		

Notes: Method: Maximum Likelihood Estimation with Robust Standard Errors and Satorra-Bentler (S-B) correction. Values below the specified thresholds are highlighted in red. Fornell-Lacker criterion not met (AVE > squared correlation of the latent construct with the discriminant construct). N = 1, 958 individual voice ratings; *** p < .001.
Global Model Fit Indices: S-Bχ^2 (53) = 1,165.865 (p < .001); χ^2/df = 41.1; Standardized Root Mean Square Residual (SRMR) = .10; Root Mean Square Error of Approximation (RMSEA) = .10; Comparative Fit Index (CFI) = .87; Tucker-Lewis Index (TLI) = .84; Goodness-of-Fit Index (GFI) = .90; Adjusted Goodness of Fit Index (AGFI) = .85; Normed Fit Index (NFI) = .87.
Abbreviations: CR = Composite Reliability; AVE = Average Variance Extracted.

Appendix H. Scree Plot of Unrotated Factors



Notes: The red dotted line indicates the "elbow" of the plot, which represents the threshold for retaining the initial factors extracted from the observed variables that maximize the variance accounted for (Eigenvalues > 1.0).

Appendix I. Exploratory Factor Analysis (Final Result)

Factor/ Item	Standardized Factor Loading (|≥.5|)	Indicator Reliability (≥.4)	Communality (≥.4)	Item-Total Correlation (≥.4)	Cronbach's Alpha (>.7)	Eigenvalue (>1.0)
Factor 1					.83	3.31
cool	.74	.54	.62	.76		
exciting	.80	.63	.63	.78		
trendy	.68	.46	.50	.69		
adventurous	.67	.45	.53	.68		
Factor 2					.90	2.22
reliable	.65	.42	.59	.74		
friendly	.51	.26	.58	.68		
honest	.73	.54	.57	.74		
real	.75	.56	.50	.68		
sincere	.73	.53	.59	.75		
down-to-earth	.63	.40	.51	.69		
responsible	.67	.45	.61	.75		
stable	.62	.38	.55	.69		
Factor 3					.83	1.62
sentimental	.56	.31	.49	.66		
expresses tender feelings	.67	.45	.61	.76		
fragile	.74	.54	.43	.54		
sensitive	.71	.50	.59	.75		
tender	.67	.45	.60	.77		
Factor 4					.68	1.53
corporate	.65	.42	.52	.62		
upper class	.57	.32	.53	.62		
Factor 5					.67	1.35
rugged	.73	.53	.51	.61		
tough	.66	.44	.53	.61		

Notes: Method: Principal Axis Factoring (PAF). **Rotation**: Oblimin with Kaiser Normalization. n = 979 individual voice ratings. **Global Model Fit Indices**: Root Mean Square Error of Approximation (RMSEA) = .04; Standardized Root Mean Square Residual (SRMR) = .02; Tucker-Lewis Index (TLI) = .97

Appendix J. Assessment of Univariate & Multivariate Normality

Factor/ Item	Min	Max	Skewness (|<2|)	p-value (>.05)	Kurtosis (|<7|)	p-value (>.05)
Factor 1						
cool	1.0	5.0	.39	.001	-.89	.001
exciting	1.0	5.0	.78	.001	-.44	.001
trendy	1.0	5.0	.66	.001	-.61	.001
adventurous	1.0	5.0	.48	.001	-.84	.001
Factor 2						
reliable	1.0	5.0	-.23	.001	-.92	.001
friendly	1.0	5.0	-.40	.001	-.71	.001
honest	1.0	5.0	-.62	.001	-.27	.001
real	1.0	5.0	-.87	.001	.12	.001
sincere	1.0	5.0	-.57	.001	-.40	.001
down-to-earth	1.0	5.0	-.49	.001	-.66	.001
responsible	1.0	5.0	-.40	.001	-.78	.001
stable	1.0	5.0	-.32	.001	-.86	.001
Factor 3						
sentimental	1.0	5.0	.27	.001	-1.00	.001
expresses tender feelings	1.0	5.0	.08	.001	-1.19	.001
fragile	1.0	5.0	.83	.001	-.45	.001
sensitive	1.0	5.0	.16	.001	-1.10	.001
tender	1.0	5.0	.21	.001	-.98	.001
Factor 4						
corporate	1.0	5.0	.61	.001	-.85	.001
upper class	1.0	5.0	.40	.001	-.85	.001
Factor 5						
rugged	1.0	5.0	.75	.001	-.61	.001
tough	1.0	5.0	.61	.001	-.74	.001
Multivariate			4,894.44	.001	52.61	.001

Notes: p-values < .05 indicate that univariate and multivariate normal distribution did not hold for the present data, although skewness and kurtosis values are within normal distribution thresholds of |<2| and |<7|, respectively (West et al., 1995). Multivariate normality was assessed using Mardia's coefficient (West et al., 1995).

Appendix K. Pairwise Correlation Coefficients Between the Acoustical Measures

	Speaking Rate	Average Silent Pause Duration	Silent Pause Frequency per Minute	Articulation Rate	Intensity Variability	Mean f0	SD f0	Range f0	h1-h2	CPPS	Jitter	Shimmer	HNR	Spectral Slope	Spectral Tilt
Speaking Rate	1.00														
Average Silent Pause Duration	-.44	1.00													
Silent Pause Frequency per Minute	-.52	.57	1.00												
Articulation Rate	.23	-.20	-.04	1.00											
Intensity Variability	-.40	.51	.55	-.03	1.00										
f0 Mean	.56	-.18	-.20	.04	-.15	1.00									
f0 SD	.28	-.17	-.24	-.07	-.32	**.75**	1.00								
f0 Range	.07	-.23	-.22	.04	-.33	.43	**.85**	1.00							
h1-h2	-.22	.28	.46	.02	.29	.09	.02	-.20	1.00						
CPPS	.30	.03	-.20	-.53	.22	.23	-.01	-.18	-.17	1.00					
Jitter	-.38	.06	.09	.47	-.07	-.57	-.24	.08	-.03	**-.72**	1.00				
Shimmer	-.12	-.05	.07	.43	-.24	-.37	-.01	.15	-.05	-.66	**.76**	1.00			
HNR	.26	.03	.00	-.28	.20	.65	.27	.02	.18	.48	**-.74**	**-.79**	1.00		
Spectral Slope	.39	-.07	-.19	-.15	-.06	.39	.10	-.07	-.13	.56	-.47	-.34	.18	1.00	
Spectral Tilt	-.39	.04	.15	.17	-.05	-.35	-.12	.02	.33	-.68	.58	.42	-.20	-.77	1.00

Notes: Correlation coefficients greater than |.8| indicate a strong pairwise correlation and are highlighted in red. Additionally, correlation coefficients greater than |.7| indicate a medium pairwise correlation and are highlighted in green.

Abbreviations: CPPS = smoothed cepstral peak prominence; f0 = fundamental frequency; HNR = harmonics-to-noise ratio; SD = standard deviation.

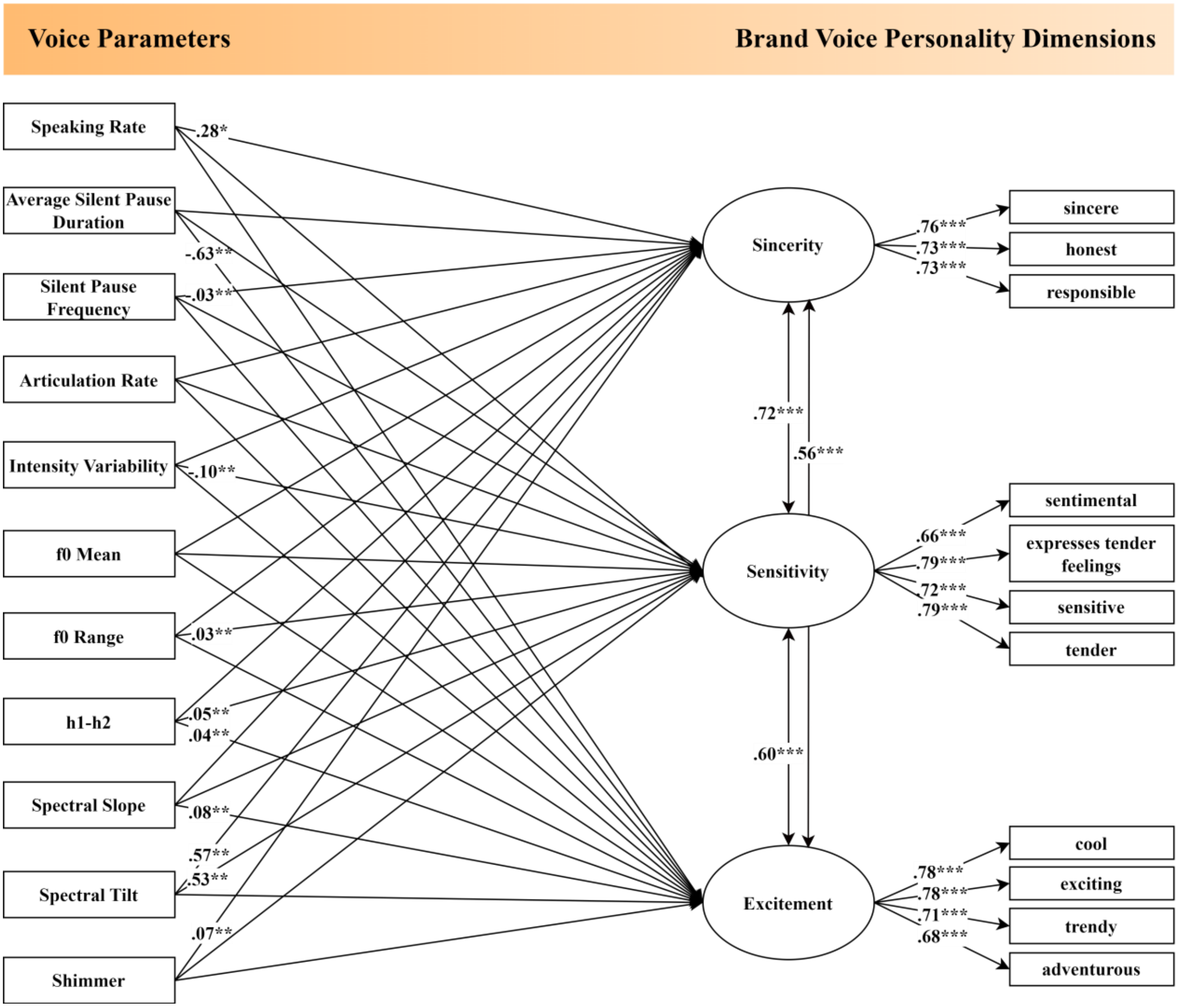
Appendix L. Consecutive Multicollinearity Assessments

Acoustical Measures	1st Assessment			2nd Assessment			Final Assessment		
	R^2	Tolerance > .2	VIF < 5	R^2	Tolerance > .2	VIF < 5	R^2	Tolerance > .2	VIF < 5
Speaking Rate	.75	.25	4.00	.75	.25	4.00	.69	.31	3.23
Silent Pause Duration	.47	.53	1.89	.44	.56	1.79	.44	.56	1.79
Silent Pause Frequency/ min	.62	.38	2.63	.60	.40	2.50	.57	.43	2.33
Articulation Rate	.67	.33	3.03	.61	.39	2.56	.36	.64	1.56
Intensity Variability	.64	.36	2.78	.64	.36	2.78	.51	.49	2.04
f0 Mean	.80	.20	5.00	.74	.26	3.85	.72	.28	3.57
f0 Range	.60	.40	2.50	.56	.44	2.27	.56	.44	2.27
h1-h2	.48	.52	1.92	.47	.53	1.89	.47	.53	1.89
Spectral Slope	.68	.32	3.13	.67	.33	3.03	.66	.34	2.94
Spectral Tilt	.79	.21	4.76	.77	.23	4.35	.73	.27	3.70
Shimmer	.68	.32	3.13	.62	.38	2.63	.55	.45	2.22
CPPS	.84	.16	**6.25**	.83	.17	**5.88**			
Jitter	.85	.15	**6.67**						

Notes: VIF values exceeding 5 indicate high (multi)collinearity and are highlighted in red. VIF = $1/(1- R^2)$; Tolerance = $1 – R^2$.
Abbreviations: CPPS = smoothed cepstral peak prominence; f0 = fundamental frequency; SD = standard deviation; VIF = variance inflation factor.

Appendix M. American Brand Voice Personality Models (BVP-Models)

A. American BVP-Model for Female Voices



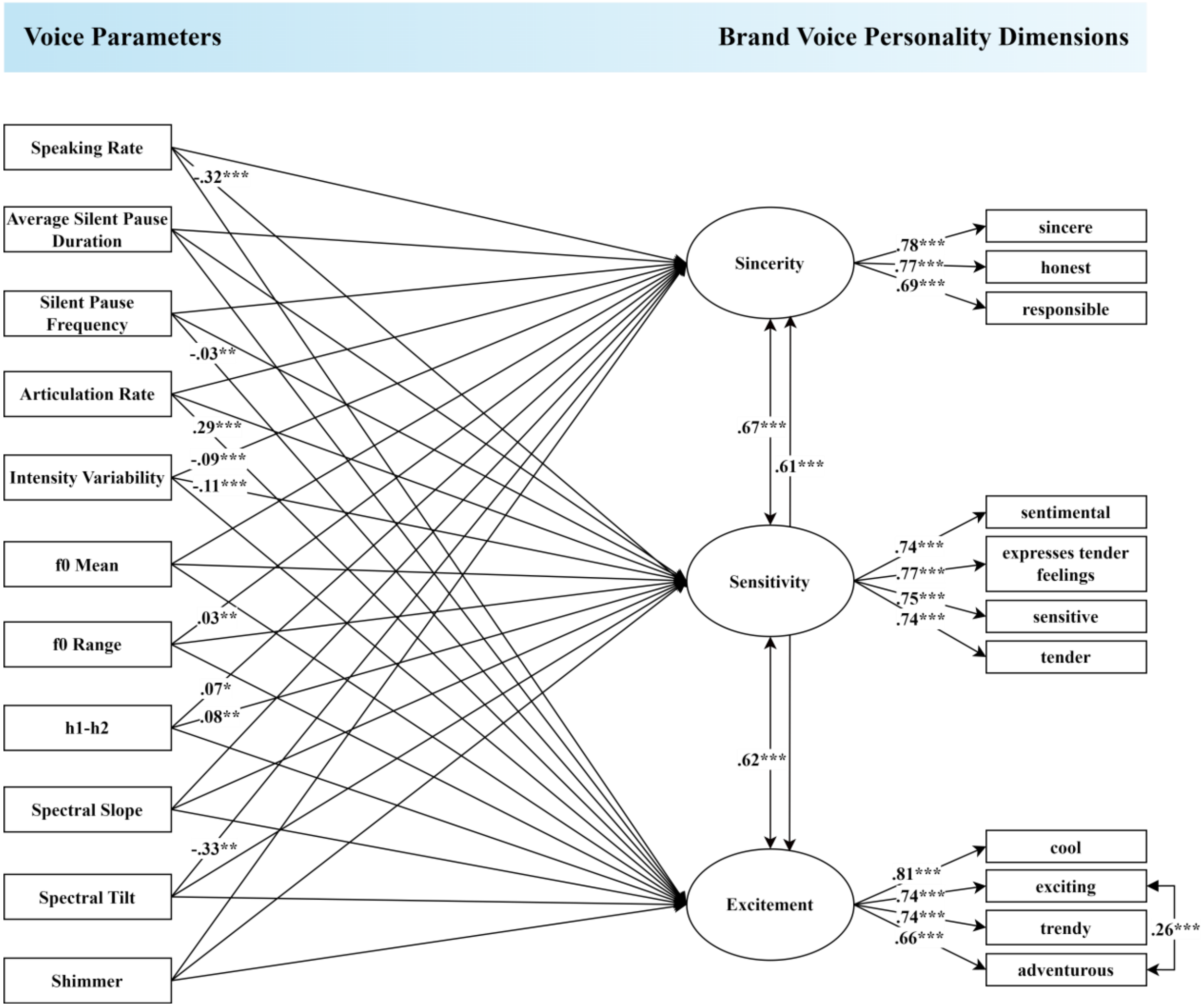
Notes: This figure is simplified to provide a better overview of the model: (1) only the significant paths are displayed; (2) socio-demographic and survey-related data are included as covariates in the analysis but are not shown in this model but displayed below. *p < .1, **p < .05, ***p < .001.

Significant Covariates:
Listener's Gender → Sincerity: β = .17**
Device Used → Sincerity: β = .15*

Listener's Gender → Sensitivity: β = .24**
Listener's Income → Sensitivity: β = .06**

Listener's Gender → Excitement: β = .26***
Listener's Age → Excitement: β = -.08***
Listener's Income → Excitement: β = .07**

B. American BVP-Model for Male Voices



Notes: This figure is simplified to provide a better overview of the model: (1) only the significant paths are displayed; (2) socio-demographic and survey-related data are included as covariates in the analysis but are not shown in this model but displayed below. *p < .1, **p < .05, ***p < .001.

Significant Covariates:
Listener's Gender → Sensitivity: β = .16**
Listener's Age → Sensitivity: β = -.08***
Listener's Income → Sensitivity: β = .06*
Background Noise → Sensitivity: β = .20**
Device Used → Sensitivity: β = .17*

Listener's Gender → Excitement: β = .19**
Listener's Age → Excitement: β = -.09***

Appendix N. Final Confirmatory Factor Analysis Result for German Brand Voice Personality Scale

Construct/ Item	Standardized Factor Loading (≥|.7|)	Item-Total Correlation (≥.4)	CR (>.7)	AVE (>.5)	Fornell/Lacker Criterion
Sincerity (α=.82)					
sincere	.82***	.77	.82	.60	Yes
honest	.75***	.73			
reliable	.75***	.72			
Sensitivity (α=.88)					
sensitive	.87***	.83	.88	.66	Yes
sentimental (Aaker, 1997)	.86***	.84			
fragile	.79***	.74			
smooth	.72***	.72			
Excitement (α=.83)					
adventurous	.76***	.74	.83	.56	Yes
exciting	.75***	.73			
spirited	.75***	.73			
daring (Grohman, 2009)	.72***	.71			

Notes: Method: Maximum likelihood estimation with robust standard errors and Satorra-Bentler (S-B) correction. n = 1,975 individual voice ratings; ***p < .001. Fornell/Lacker criterion: AVE > squared correlations between the constructs.
Abbreviations: CR = Composite Reliability; AVE = Average Variance Extracted.
Global Model Fit Indices: S-Bχ^2 (41) = 146.971 (p < .001); χ^2/df = 4.3; Standardized Root Mean Square Residual (SRMR) = .03; Root Mean Square Error of Approximation (RMSEA) = .04; Comparative Fit Index (CFI) = .99; Tucker-Lewis Index (TLI) = .98; Goodness-of-Fit Index (GFI) = .98; Adjusted Goodness of Fit Index (AGFI) = .97; Normed Fit Index (NFI) = .98.

Appendix O. American & German Estimated Regression Coefficients of the Structural Equation Models for Sincerity

Relationship	Path	American English						German					
		Female Voices $R^2 = .071$			Male Voices $R^2 = .062$			Female Voices $R^2 = .036$			Male Voices $R^2 = .025$		
		β	S.E.	z-value	β	S.E.	z-value	β	S.E.	z-value	β	S.E.	z-value
Acoustical Measures → Sincerity	Speaking Rate → Sincerity	.28	.14	1.69*	.06	.09	.67	.12	.06	2.67**	.12	.09	2.22**
	Av. Pause Duration → Sincerity	.28	.19	1.23	.04	.16	.26	.03	.08	.56	-.03	.04	-1.08
	Silent Pause Frequency → Sincerity	-.03	.01	-2.97**	.00	.01	.16	-	-	-	-	-	-
	Articulation Rate → Sincerity	-.22	.12	-1.58	-.01	.08	-.06	-.05	.05	-1.37	-.06	.10	-1.22
	Intensity Variability → Sincerity	.05	.04	1.04	-.09	.02	-3.58***	.00	.03	.04	.03	.03	.69
	f0 Mean → Sincerity	-.02	.02	-1.11	.00	.03	-.10	-.02	.01	-.71	-.01	.02	-.33
	f0 Range → Sincerity	.00	.01	-.36	.03	.01	2.30**	.01	.01	.24	.02	.01	.76
	h1-h2 → Sincerity	.01	.02	.45	.07	.03	2.10**	.12	.01	3.56***	.01	.01	.19
	Spectral Slope → Sincerity	-.02	.03	-.73	-.02	.02	-.74	.09	.01	2.12**	.07	.01	1.84*
	Spectral Tilt → Sincerity	.57	.15	3.17**	-.33	.13	-2.39**	-.05	.04	-1.40	.07	.04	2.18**
	Shimmer → Sincerity	.07	.03	1.95**	.00	.03	-.02	-.03	.03	-.53	-.04	.03	-.84
	CPPS → Sincerity	-	-	-	-	-	-	-.12	.03	-2.31**	-.09	.03	-1.46
Covariates → Sincerity	Listener's Gender[a] → Sincerity	.17	.06	2.37**	.11	.07	1.41	.01	.04	.37	-.08	.05	-3.02**
	Listener's Age → Sincerity	.02	.02	.92	-.04	.02	-1.39	-.02	.01	-.55	-.02	.02	-.81
	Listener's Education → Sincerity	.00	.02	-.16	-.02	.02	-.80	.02	.01	.71	.00	.01	.00
	Listener's Profession → Sincerity	.00	.02	1.03	.01	.02	.56	-.01	.01	-.38	.03	.01	.99
	Listener's Income → Sincerity	.04	.03	.93	.04	.03	1.35	-.04	.03	-1.35	-.02	.03	-.84
	Background Noise → Sincerity	.09	.07	1.11	.15	.09	1.50	.02	.06	.58	.02	.06	.84
	Device Used[b] → Sincerity	.15	.08	1.66*	.09	.08	.97	-.02	.05	-.78	-.01	.05	-.53

Notes: Estimates represent standardized path coefficients. Significant paths have been highlighted in blue to improve orientation. The darker the color, the higher the significance level. [a] dummy coded: "0" = female, "1" = male; [b] dummy coded: "0" = speaker, "1" = headphones; *p < .1; ** p < .05; ***p < .001.
Abbreviations: av. = average; S.E. = standard errors.

Appendix P. American & German Estimated Regression Coefficients of the Structural Equation Models for Sensitivity

		American English						German					
		Female Voices $R^2 = .124$			Male Voices $R^2 = .098$			Female Voices $R^2 = .080$			Male Voices $R^2 = .044$		
Relationship	**Path**	**β**	**S.E.**	**z-value**	**β**	**S.E.**	**z-value**	**β**	**S.E.**	**z-value**	**β**	**S.E.**	**z-value**
Acoustical Measures → Sensitivity	Speaking Rate → Sensitivity	-.15	.13	-.97	-.32	.09	-3.41***	.07	.10	1.39	.07	.10	1.39
	Av. Pause Duration → Sensitivity	-.08	.18	-.38	.03	.15	.21	-.03	.04	-1.12*	-.03	.04	-1.12
	Silent Pause Frequency → Sensitivity	.00	.01	.40	-.01	.01	-.77	-	-	-	-	-	-
	Articulation Rate → Sensitivity	.21	.11	1.64	.13	.08	1.63	-.04	.11	-.94*	-.04	.11	-.94
	Intensity Variability → Sensitivity	-.10	.04	-2.27**	-.11	.02	-4.11***	-.07	.03	-1.61**	-.07	.03	-1.61
	f0 Mean → Sensitivity	.01	.02	.40	-.01	.03	-.23	-.01	.02	-.29*	-.01	.02	-.29
	f0 Range → Sensitivity	.03	.01	2.26**	.01	.01	.94	.03	.01	.80	.03	.01	.80
	h1-h2 → Sensitivity	.05	.02	2.63**	.08	.03	2.53**	.06	.02	1.95***	.06	.02	1.95**
	Spectral Slope → Sensitivity	.04	.03	1.20	.00	.02	-.12	.05	.01	1.46	.05	.01	1.46
	Spectral Tilt → Sensitivity	.53	.15	3.10**	-.14	.13	-1.05	.06	.04	2.09	.06	.04	2.09**
	Shimmer → Sensitivity	-.02	.03	-.61	-.04	.03	-1.16	-.09	.03	-2.03**	-.09	.03	-2.03**
	CPPS → Sensitivity	-	-	-	-	-	-	-.12	.04	-2.25***	-.12	.04	-2.25**
Covariates → Sensitivity	Listener's Gender[a] → Sensitivity	.24	.06	3.40***	.16	.07	2.24**	.10	.05	4.17***	.05	.05	1.82*
	Listener's Age → Sensitivity	-.02	.02	-.63	-.08	.02	-3.47***	.06	.02	2.36**	.00	.02	-.11
	Listener's Education → Sensitivity	-.03	.02	-1.21	.00	.02	.11	-.04	.01	-1.48	-.01	.01	-.34
	Listener's Profession → Sensitivity	.01	.01	.54	.02	.02	1.22	-.02	.01	-.57	-.01	.01	-.34
	Listener's Income → Sensitivity	.06	.02	2.11**	.06	.03	1.92*	-.06	.03	-2.18**	-.05	.03	-1.96*
	Background Noise → Sensitivity	.03	.07	.30	.20	.09	2.08**	.01	.07	.33	.03	.07	1.09
	Device Used[b] → Sensitivity	.10	.08	1.10	.17	.09	1.83*	.01	.06	.50	.01	.06	.49

Notes: Estimates represent standardized path coefficients. Significant paths have been highlighted in blue to improve orientation. The darker the color, the higher the significance level. [a] dummy coded: "0" = female, "1" = male; [b] dummy coded: "0" = speaker, "1" = headphones; *p < .1; ** p < .05; ***p < .001.
Abbreviations: av. = average; S.E. = standard errors.

Appendix Q. American & German Estimated Regression Coefficients of the Structural Equation Models for Excitement

		American English						German					
		Female Voices $R^2 = .107$			Male Voices $R^2 = .104$			Female Voices $R^2 = .101$			Male Voices $R^2 = .099$		
Relationship	**Path**	**β**	**S.E.**	**z-value**	**β**	**S.E.**	**z-value**	**β**	**S.E.**	**z-value**	**β**	**S.E.**	**z-value**
Acoustical Measures → Excitement	Speaking Rate → Excitement	-.01	.15	-.06	-.11	.10	-1.18	.20	.05	4.46***	.07	.08	1.31
	Av. Pause Duration → Excitement	-.63	.23	-2.70**	.13	.17	.83	.09	.07	1.98**	.05	.04	1.66*
	Silent Pause Frequency → Excitement	-.01	.01	-1.11	-.03	.01	-2.94**	-	-	-	-	-	-
	Articulation Rate → Excitement	.18	.13	1.30	.29	.09	3.54***	.01	.05	.34	.06	.09	1.16
	Intensity Variability → Excitement	-.05	.04	-1.24	-.01	.03	-.20	.04	.02	.89	-.03	.02	-.72
	f0 Mean → Excitement	.01	.02	.41	-.04	.03	-1.42	-.02	.01	-.50	.01	.01	.28
	f0 Range → Excitement	.00	.01	.33	.01	.01	.59	.01	.01	.23	-.03	.01	-1.10
	h1-h2 → Excitement	.04	.02	1.97**	-.01	.03	-.36	.07	.01	2.29**	-.02	.01	-.71
	Spectral Slope → Excitement	.08	.03	2.66**	.04	.03	1.49	.14	.01	3.30***	.14	.01	3.56***
	Spectral Tilt → Excitement	.19	.17	1.14	-.13	.15	-.93	-.01	.03	-.36	.00	.04	.02
	Shimmer → Excitement	.00	.03	-.02	.00	.03	-.06	-.09	.02	-2.01**	-.07	.02	-1.65*
	CPPS → Excitement	-	-	-	-	-	-	-.08	.03	-1.62	-.15	.03	-2.63**
Covariates → Excitement	Listener's Gender[a] → Excitement	.26	.07	3.54***	.19	.08	2.56**	.11	.04	4.26***	.08	.04	3.02**
	Listener's Age → Excitement	-.08	.02	-3.74***	-.09	.02	-3.99***	-.23	.01	-8.38***	-.22	.01	-8.57***
	Listener's Education → Excitement	-.03	.02	-1.40	-.01	.02	-.25	-.08	.01	-2.85**	-.07	.01	-2.63**
	Listener's Profession → Excitement	.02	.02	.99	-.01	.02	-.56	-.03	.01	-.93	.03	.01	1.18
	Listener's Income → Excitement	.07	.03	2.52**	.03	.03	.91	.00	.02	-.04	.00	.02	.02
	Background Noise → Excitement	-.06	.09	-.66	-.02	.11	-.21	-.01	.05	-.33	.03	.05	1.30
	Device Used[b] → Excitement	.07	.09	.75	.06	.09	.68	.04	.05	1.32	.04	.05	1.45

Notes: Estimates represent standardized path coefficients. Significant paths have been highlighted in blue to improve orientation. The darker the color, the higher the significance level. [a] dummy coded: "0" = female, "1" = male; [b] dummy coded: "0" = speaker, "1" = headphones; *p < .1; ** p < .05; ***p < .001.
Abbreviations: av. = average; S.E. = standard errors.

Statutory Declaration

I hereby confirm that I have written this dissertation without unauthorized assistance and that no sources or aids other than those specified have been used. I further declare that the passages taken literally or in terms of content from the works used have been identified as such. I agree that the electronic version of the thesis is allowed to be checked in case of accusation of plagiarism.

Eidesstattliche Erklärung

Hiermit versichere ich an Eides Statt, dass die Dissertation ohne unerlaubte Hilfe angefertigt und keine anderen, als die angegebenen Quellen und Hilfsmittel benutzt wurden. Ich erkläre ferner, dass die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht wurden. Eine Überprüfung der Dissertation mit qualifizierter Software im Rahmen der Untersuchung von Plagiatsvorwürfen ist gestattet.

City and Date

Signature