Can you hear my personality? A conceptualization of a brand voice based on brand personality

Olga Bosak
University of Bremen
Maik Eisenbeiss
University of Bremen
Kristina Klein
University of Bremen

Acknowledgements:
This project has received funding from the Central Research Development Fund (CRDF) of the University of Bremen.

Cite as:
Bosak Olga, Eisenbeiss Maik, Klein Kristina (2023), Can you hear my personality? A conceptualization of a brand voice based on brand personality. *Proceedings of the European Marketing Academy*, 52nd, (114332)

Paper from the 52nd Annual EMAC Conference, Odense/Denmark, May 23-26, 2023

EMAC
EUROPEAN MARKETING ACADEMY

Can you hear my personality? A conceptualization of a brand voice based on brand personality

Abstract:

Companies advertising and interacting with their customers via smart speakers communicate predominantly auditorily. In this type of communication, the voice plays a major role, as the personality of the brand is perceived exclusively through the voice. But how does a voice have to sound in order for a desired brand personality to be perceived by users? Using an exploratory approach, we show that brand personalities of confidence, sensitivity, and excitement can be perceived through voice alone. We also determine which combinations of voice features are crucial for the perception of these brand personalities. With our findings, we provide guidance to marketers, voice interface architects, and UX designers on how to translate a brand personality into a voice so that voice assistant users can auditorily perceive the desired personality of a brand.

Keywords: brand voice, brand personality, voice assistants

Track: *Product and Brand Management*

1. Introduction

Consumers mainly interact with technological devices by clicking or typing. However, consumers also increasingly use their voice to search for information, shop, get directions or make reservations (Melzner, Bonezzi, and Meyvis, 2022). Voice-based conversational agents, such as Amazon's Alexa, Google Assistant or Apple's Siri, offer a human-like interaction with their users (Sciuto, Saini, Forlizzi, and Hong, 2018). Global sales of smart speakers with voice assistants installed increased from 31.7 million units in 2017 to 145.8 million units in 2020, a 360% increase within four years (Paxton, 2021), demonstrating the rapid diffusion of voice technology. According to forecasts, global sales of smart speakers will continue to grow steadily over the next 5 years (Munster & Thompson, 2019).

For marketers, the emergence of voice assistants represents an opportunity as a new key communication channel and advertising medium (Lee & Cho, 2020). Regarding the latter, the largest providers of smart speakers or voice assistants, Amazon and Google, offer indirect and direct ways to advertise on their devices. Users of smart speakers can listen to radio or Spotify songs, and podcasts with included advertising from these platforms. Furthermore, users of Amazon's Alexa can directly receive personalized recommendations for products based on their previous shopping behavior (Hardesty, 2019). More importantly though, companies can develop their own voice applications for these smart speakers to provide entertainment, information or services to users (Amazon, 2022). In this context, the question arises how a company should sound like. With smart speakers, the interaction with users is limited to purely auditory communication. Thus, the voice (next to the tonality or the content provided) will communicate the personality of the brand. Consequently, marketers need to find a voice that fits their brand so that customers can perceive the brand personality in an auditory communication. However, no systematic approach how to find the "right voice" to transport a brand's personality exists so far – this study wants to close this gap.

2. Research Objective

Research in phonetics demonstrates the importance of voice and certain voice features, e.g., pitch, speaking rate, and intonation, in (human) speakers' perception (Schweinberger & Zäske, 2019). In addition to verbal content, a speaker conveys non-verbal content to the listener, such as an emotional and motivational state (Gobl & Chasaide, 2003), physiological cues (Krauss, Freyberg, and Morsella, 2002) and his/her identity and personality (Zäske, Skuk, Golle, and Schweinberger, 2020). For example, a person with a low-pitched voice is likely to be perceived as more competent, confident, and trustworthy (Oleszkiewicz, Pisanski, Lachowicz-Tabaczek, and Sorokowska, 2017; Rodero, 2013); a person with a rather fast speaking rate is likely to be perceived as more extroverted and ambitious (Addington, 1968; Brown, Giles, and Thakerar, 1985).

So far, no single comprehensive set of voice features that is essential for voice personality recognition exists. Studies that analyze effects of voice on personality perception mostly focus on the combination of only a few voice features such as pitch or speaking rate (Brown et al., 1985; Dahl, 2010). Those studies often do not consider interaction effects between individual voice features (Apple, Streeter, and Krauss, 1979). Further, most studies on associations of voice qualities (e.g., breathiness or hoarseness) with personality traits rely on ratings from experienced judges. That is, they do not measure the voice features objectively but rely on subjective ratings, which renders the results difficult to generalize (Addington, 1968).

This study builds on previous findings regarding voice and personality research in phonetics, psychology, and marketing. First, we want to find out to what extent people perceive brand personalities through voice by developing a scale to measure a brand's voice personality. Second, we investigate which combination of voice features influence specific brand personality perceptions. Importantly, we only consider objectively measurable voice features and their interaction effects. With our findings, we provide directions for marketers, voice user interface architects, and conversational UX designers on how to transcribe a brand personality into a voice, so that voice assistant users can auditorily perceive the desired personality.

3. Methodology

3.1. Voice features, vocal stimuli and brand personality measures

The voice features we use in our study reflect previous findings in psychology regarding correlations between specific personality traits and are acoustically measurable (see Table 1). The voice features can be described according to their perception by listeners and four distinct soundwave dimensions: timing, amplitude, frequency, and spectral (Hildebrand et al., 2020; Jurafsky & Martin, 2020).

Voice feature (metric)	Listener's perception	Soundwave dimension
Speaking rate (syl/s)	Fluency	Timing
Silent pause duration (s)		
Articulation rate (syl/s)	Velocity of speech	
Intensity variability (SD)	Loudness variability	Amplitude/ Intensity
Fundamental frequency mean (f0 mean; Hz)	Pitch	Frequency
Fundamental frequency range (f0 SD; st)	Pitch variability	
h1-h2 (dB)	Roughness	Spectral (Voice Quality)
Acoustic Breathiness Index (ABI)	Breathiness	
Acoustic Voice Quality Index (AVQI)	Hoarseness	

Table 1. Voice features for personality perception

Note: syl = syllable; s = second; SD = standard deviation; Hz = Hertz; st = semitones; dB = Decibel

We used voice samples from the Jena Speaker Set (JESS) as vocal stimuli (Zäske et al., 2020). JESS is a database of voice recordings from 120 female and male speakers comprising German sentences, read text, semi-spontaneous speech, syllables, and sustained vowel stimuli. For this study, we only used the semi-spontaneous speech recordings, which are descriptions of a farmyard scene. Personality traits and emotions have been shown to be best reflected when speakers speak spontaneously and have no behavioral constraints, making the semi-spontaneous recordings the ideal choice for our research purpose (Johnstone & Scherer, 2000). Of this database, we only used 96 voice samples (47 female), since 24 voices showed audible dialects from eastern and southern parts of Germany. As dialects can be associated with specific social classes, this could have affected the perception of personalities (Krauss et al., 2002). Thus, we removed those voice samples. On average, respondents rated each voice sample 41 times.

In order to objectively analyze the voice samples regarding their voice features, we used the software Praat (version 6.1.50; Boersma & Weenink, 1992-2022). Linguistics and communication sciences widely use this software tool for phonetic and speech analysis because of its user-friendly interface and a number of publicly available extensions, plugins and scripts. By using a Praat script designed by the authors, all vocal measures are calculated automatically and their numerical output is stored.

For the personality perception rating, we focused on 62 personality traits taken from the brand personality scales (BPS) of Aaker (1997), Geuens, Weijters, and de Wulf (2009), and Grohmann (2009; see Table 2). These generalizable BPS focus on consumer brands and apply to diverse product categories, countries, and cultures.

Authors (Year)	Brand personality traits
Aaker (1997)	charming, cheerful, confident, contemporary, cool, corporate, **daring**, **down-to-earth**, exciting, family-oriented, feminine, friendly, glamorous, good-looking, hard-working, honest, imaginative, independent, intelligent, leader, masculine, original, outdoorsy, real, reliable, rugged, secure, **sentimental**, sincere, small-town, smooth, spirited, successful, technical, tough, trendy, unique, upper, class, up-to-date, western, wholesome, young
Geuens et al. (2009)	active, **aggressive**, bold, **down-to-earth**, dynamic, innovative, ordinary, responsible, romantic, **sentimental**, simple, stable
Grohmann (2009)	adventurous, **aggressive**, brave, **daring**, dominant, expresses tender feelings, fragile, graceful, sensitive, sturdy, sweet, tender

Table 2. Overview of used brand personality traits

Note: In total there are 66 items, but four items occur twice within the three brand personality scales (see bolded items).

3.2. Study design and sample

In an initial first step, as our approach is exploratory, we designed an online questionnaire consisting of three sections. First, participants answered demographic and filter questions (no hearing impairment, only German-speaking). Next, participants listened to and rated a first voice sample on a five-point scale (from 1 (“does not apply at all”) to 5 (“applies fully”)) as to what extent they associated the voice with the personality traits provided. Afterwards, they did the same for a second voice sample. Participants were randomly assigned to both voice samples. All participants provided informed consent before the research began. 2,123 participants completed the questionnaire (through a German online-access panel provider). We removed 123 participants with insufficient effort (missing answers, not passing attention check, not playing the voice sample, response patterns and speeders) to answer the questionnaire (Huang, Curran, Keeney, Poposki, and DeShon, 2012; Leiner, 2019), resulting in 2,000 participants with 3,945 individual voice ratings (1,945 participants with two and 55 participants with one voice ratings; M_{age} = 51, 53% female).

4. Results

4.1. Development of a brand voice personality scale

Our first research objective was to find out to what extent participants perceive brand personalities through voice. We thus develop the brand voice personality scale (BVPS), indicating which brand personality traits and dimensions can be perceived through voice alone. First, we split the sample of 3,945 voice ratings into two equivalent subsamples ($N_{subsample1}$ = 1,973; $N_{subsample2}$ = 1,972) according to the Solomon method (Lorenzo-Seva, 2021). Second, we conducted an explorative factor analysis (EFA) with the 62 personality traits with subsample 1 in order to investigate the dimensionality of the BVPS. The EFA with an oblique rotation (oblimin) with Kaiser normalization resulted in a three-factorial model with 37 personality traits (Hair, Howard, and Nitzl, 2020; Kaiser & Rice, 1974). All constructs showed factor loadings >.50, eigenvalues >1, and Cronbach's alpha values >.80 suggesting high internal consistency (DeVellis & Thorpe, 2021). Thirdly, to validate the developed BVPS, we conducted confirmatory factor analyses (CFA) with subsample 2, resulting in a reflective first-order model with three personality dimensions (confidence, sensitivity, and excitement), incorporating 24 personality traits (see Table 3).

Construct/ Item	**Standardized Factor Loadings**	**Composite Reliability**	**Average Variance Extracted**
Confidence		.92	.52
contemporary	.72***		
wholesome	.69***		
reliable	.73***		
confident	.77***		
independent	.70***		
sincere	.66***		
secure	.71***		
active	.73***		
stable	.67***		
responsible	.74***		
intelligent	.76***		
Sensitivity		.92	.57
sensitive	.81***		
sentimental (from Geuens et al., 2009)	.66***		
sweet	.74***		
smooth	.71***		
romantic	.71***		
sentimental (from Aaker, 1997)	.84***		
fragile	.81***		
expresses tender feelings	.78***		
tender	.73***		
Excitement		.83	.55
spirited	.73***		
adventurous	.77***		
daring (from Grohman, 2009)	.73***		
exciting	.74***		

Table 3. Brand voice personality scale (items translated to English)

Note: Method: maximum likelihood estimation with robust standard errors and Satorra-Bentler (S-B) correction.
S-B$\chi 2$ = 2,090.741 (df = 249, p = .00); RMSEA = .061; SRMR = .053; AGFI = .888; CFI = .933; TLI = .925. *** p < .001

4.2. Perception of brand personalities through voice

Our second research objective was to find out which combination of voice features influence perceptions of brand personalities. To analyze the data, we used a multigroup structural equation model (SEM) with a group code approach. This model allows for the estimation of group-specific effects in the context of a SEM. In our case, we used the feminine vs. masculine voices to form two groups. The structural model shows a moderate model fit (S-Bχ2 = 9,685.284($_{906 \text{ df}}$), $p < .001$; RMSEA = .075; SRMR = .066; CFI = .861), which is not surprising due to the exploratory nature of our study.

Figure 1 depicts the structural model for a) feminine voices and b) masculine voices.

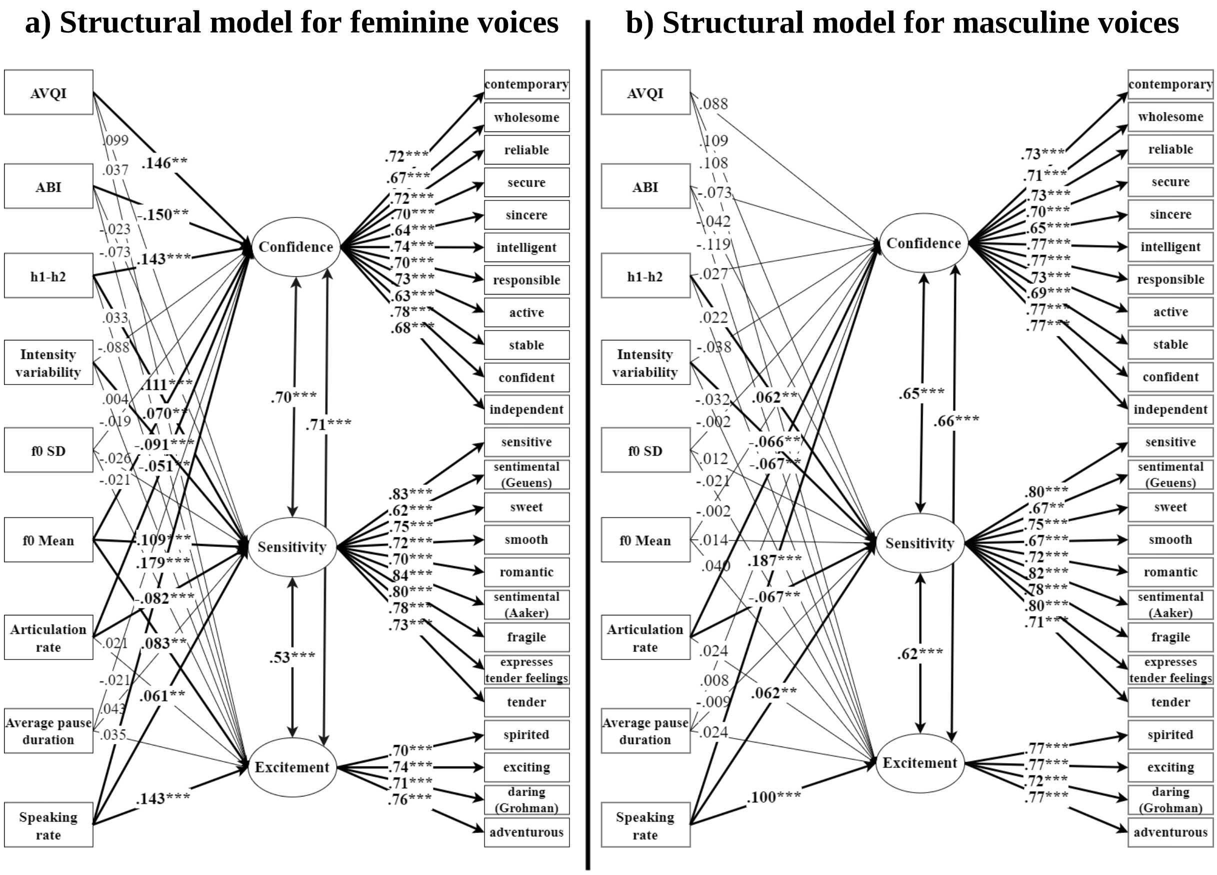

Figure 1. Structural models

Note: For presentation purposes significant correlation paths are shown in bold and associated factor loadings are centered. SD = standard deviation; ABI = Acoustic Breathiness Index; AVQI = Acoustic Voice Quality Index; ** p < .05; *** p < 0.001

For feminine voices, the structural model shows positive correlations for confidence with speaking rate (β = .179, z = 6.309, $p < .001$), f0 mean (β = .070, z = 2.645, $p < .05$), h1-h2 (β = .143, z = 4.989, $p < .001$), and AVQI (β = .146, z = 2.565, $p < .05$) and negative correlations with articulation rate (β = -.051, z = -2.203, $p < .05$) and ABI (β = -.150, z = -2.851, $p < .05$). For the sensitivity, we identify significant positive correlations with speaking rate ($\beta = .061$, $z = 2.179$, $p < .05$), f0 mean ($\beta = .109$, $z = 4.208$, $p < .001$), and h1-h2 ($\beta = .111$, $z = 3.921$, $p < .001$), and negative correlations with articulation rate ($\beta = -.082$, $z = -3.513$, $p < .001$) and intensity variability ($\beta = -.091$, $z = -3.412$, $p < .001$). For excitement, the results indicate significant positive correlations with speaking rate ($\beta = .143$, $z = 4.484$, $p < .001$) and f0 mean ($\beta = .083$, $z = 2.987$, $p < .05$).

Based on the results, a feminine brand that speaks more fluent but not fast with a high pitch transmits a confident personality. Furthermore, it is advantageous if the voice has a hoarseness and roughness (as indicated by AVQI and h1-h2, respectively), but is not breathy (as indicated by ABI). For the perception of a sensitive personality, female brands should have a high-pitched rough voice and speak fluently but not fast and with low intonation (as indicated by intensity variability). When a feminine brand speaks fluently with a high-pitched voice, perceptions relate to an excited brand personality.

For masculine voices, confidence shows a positive correlation with speaking rate ($\beta = .187$, $z = 6.327$, $p < .001$) and a negative correlation with articulation rate ($\beta = -.067$, $z = -2.330$, $p < .05$). For sensitivity, we identify significant positive correlations with speaking rate ($\beta = .062$, $z = 2.140$, $p < .05$) and h1-h2 ($\beta = .062$, $z = 2.366$, $p < .05$), and negative correlations with articulation rate ($\beta = -.067$, $z = -2.392$, $p < .05$) and intensity variability ($\beta = -.066$, $z = -2.418$, $p < .05$). For excitement, a significant positive correlation with speaking rate exists ($\beta = .100$, $z = 3.171$, $p < .001$).

Thus, for masculine brands, the combination of two timing voice features, namely a fluent but not fast speaking style, transports a confident personality. For the perception of a masculine sensitive personality, brands should have a rough voice and speak fluently with low intonation but not fast. Masculine excited brand personality is perceived through a fluent speaking style.

5. Conclusion

First, our results suggest that brand personalities from well-established BPS cannot fully be perceived through voice. We identify three personality dimensions (confidence, sensitivity, and excitement; based on the BVPS) that should be used to search for a suitable brand voice. Further, marketers should determine the gender of the brand before developing the personality, as the gender of the voice determines the exact combination of voice features leading to the perception of a specific personality.

When comparing the perception of confidence between both genders, it is apparent that the feminine voice must have more vocal facets than a masculine one, since the voice qualities hoarseness, roughness, and breathiness play an important role. In contrast, the sensitivity and excitement brand personalities differ between genders only with respect to one voice feature, and that is the fundamental frequency (f0). For the perception of both brand personalities, it is advantageous for feminine brands to have a high-pitched voice, although this does not seem to have any influence with masculine brand voices.

By using German voice samples and brand personality traits, this study focusses on one language area only. In order to test the cross-linguistic validity of the results, we plan a replication of the study in English-speaking and other countries. Moreover, we will create own voice samples with the voice features that make up specific brand personalities and test whether those are then really perceived as intended. This approach will provide a solid grounding for our proposed BVPS.

References

Aaker, J. L. (1997). Dimensions of Brand Personality. *Journal of Marketing Research*, *34*(3), 347–356.

Addington, D. W. (1968). The relationship of selected vocal characteristics to personality perception. *Speech Monographs*, *35*(4), 492–503.

Amazon. (2022). *What Are Alexa Skills?* Retrieved from https://developer.amazon.com/en-US/alexa/alexa-skills-kit. (Last accessed: December 05, 2022).

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.

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

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.

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.

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications: fifth edition*. SAGE Publications.

Geuens, M., Weijters, B., & Wulf, K. de (2009). A new measure of brand personality. *International Journal of Research in Marketing*, *26*(2), 97–107.

Gobl, C., & Chasaide, A. N. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech Communication*, *40*(1-2), 189–212.

Grohmann, B. (2009). Gender Dimensions of Brand Personality. *Journal of Marketing Research*, *46*(1), 105–119.

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.

Hardesty, L. (November 22, 2019). *The history of Amazon's recommendation algorithm. Collaborative filtering and beyond.* Retrieved from https://www.amazon.science/the-history-of-amazons-recommendation-algorithm. (Last accessed: December 05, 2022).

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.

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.

Johnstone, t., & Scherer, K. R. (2000). Vocal communication of emotion. In Lewis M. & Haviland J. (eds.), *Handbook of Emotion: second edition.* New York: Guilford.

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

Kaiser, H. F., & Rice, J. (1974). Little Jiffy, Mark IV. *Educational and Psychological Measurement*, *34*(1), 111–117.

Krauss, R. M., Freyberg, R., & Morsella, E. (2002). Inferring speakers' physical attributes from their voices. *Journal of Experimental Social Psychology*, *38*(6), 618–625.

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.

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.

Lorenzo-Seva, U. (2021). Solomon: A method for splitting a sample into equivalent subsamples in factor analysis. *Behavior Research Methods*, 1–13.

Melzner, J., Bonezzi, A., & Meyvis, T. (2022). EXPRESS: Information Disclosure in the Era of Voice Technology. *Journal of Marketing.* Advance online publication. https://doi.org/10.1177/00222429221138286

Munster, G., & Thompson, W. (June 13, 2019). *Smart Speaker Macro-Model Update*. Retrieved from https://loupfunds.com/smart-speaker-market-share-update/. (Last accessed: December 05, 2022).

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.

Paxton, M. (2021, November 4). *Alexa, tell me about the smart speaker market in 2021*. Retrieved from https://www.spglobal.com/marketintelligence/en/news-insights/blog/alexa-tell-me-about-the-smart-speaker-market-in-2021. (Last accessed: December 05, 2022).

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

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.

Sciuto, A., Saini, A., Forlizzi, J., & Hong, J. I. (2018). "Hey Alexa, What's Up?". In T. Cerratto-Pargman (ed.) *Proceedings of the 2018 Designing Interactive Systems Conference.* ACM.

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.