International Journal of Organizational Leadership 13(2024) 432-450

INTERNATIONAL JOURNAL OF ORGANIZATIONAL LEADERSHIP

WWW.CIKD.CA

journal homepage: https://www.ijol.cikd.ca

How Speech Prosody Affects Leadership Perceptions: A Machine Learning Approach

Chia-Chun Hsu[1], Jarek Krajewski[2], Jörg Felfe[3]

[1]Institute of Safety Technology, University of Wuppertal, Germany
[2]Institute of Experimental Psychophysiology, Germany
[3,1]Department of Work, Organizational, and Business Psychology, Helmut Schmidt University/University of the Federal Armed Forces, Hamburg, Germany

Keywords:
Fundamental frequency, Intensity, Voicing probability, Leadership, Machine learning

Received
12 May 2024
Received in revised form
30 May 2024
Accepted
05 June 2024

***Correspondence:**
chia-chun.hsu@uni-wuppertal.de

Abstract

This paper investigates the impact of speech prosody on perceived transformational leadership within a business context. Previous works have mainly concentrated on political preference or leadership; however, the business scenario is rather limited. Notably, a few studies have pointed to the subscale of transformational leadership, yet there is scant research to examine the full-dimensions of transformational leadership. Utilizing a large sample of 122 speakers and 122 evaluators, OpenSMILE was employed to extract the prosodic features, and Waikato Environment for Knowledge Analysis was used to analyze the data. With the SMO regression as the machine learning algorithm, the results indicated that fundamental frequency, speech intensity, and voicing probability influence various dimensions, including inspirational motivation, idealized influence, intellectual stimulation, and individualized consideration. Fundamental frequency emerges as a pivotal predictor of highlighting the importance of conveying leadership qualities across multiple subscales of transformational leadership. The findings also suggested practical implications for further leadership development by vocal cues. Nevertheless, limitations include the predominantly student sample and the specific context, warranting further research on diverse populations and backgrounds. Future research could explore cross-cultural variations in leadership perceptions and extend analyses to transactional leadership. Overall, this research enriches our understanding of speech prosody in shaping perceptions of leadership and provides insights into the aspects of theory and practice.

Speech communication is central to human life; it is vital for survival and social communication (Grossmann & Freiderici, 2012). Communication is essentially an exchanging of signals, both verbal (e.g., speech, announcement, and face-to-face interaction) and non-verbal (e.g., body language, facial expression, gesture, and voice), between two or more individuals. From this point on, speakers are viewed as powerful or dominant whether listeners perceive the signals they send as powerful or dominant (Bradac et al., 1994). From a theoretical view, both verbal and non-verbal communications have auditory characteristics: the verbal voice corresponds to the message content, and the non-verbal one specifies how the message is communicated (Hargie, 2011). Past research has shown that social status (e.g., hierarchical relationships) is most commonly communicated by non-verbal rather than verbal signals (Mignault & Chaudhuri, 2003). More precisely, when verbal and non-verbal cues of the rank are present, non-verbal cues have a higher chance to influence participants' judgment than verbal ones (Jacob et al., 2012).

Hall et al. (2005) have found that nonverbal communication is related to power. They demonstrate a clear correlation between the use of non-verbal activity by a person and the influence they receive in encounters. Our views of others are focused not only on what people say but also on the vocal cues (Sporer & Schwandt, 2006). While several verbal and non-verbal cues are possible, the authors focus on the discussion of speech parameters in this research: the nonverbal features of a voice that give rise to differing impressions of speakers. As Scherer (1982) noted, “During evolution, language and speech were superimposed on a primitive, analog vocal signaling system (p. 138)”. Thus, owing to its basic origins, speech features may be more important than linguistic material or other less primal non-verbal signals in developing social judgments (Tusing & Dillard, 2000).

Voice is not only a kind of biosignal of humans but also related to perception. It is the counterpart of speech production aiming to decode sounds and contain the message. The human voice that transmits information in daily social experiences has also been called an "auditory face" (Belin et al., 2004) because it offers valuable social information about the speaker’s identity (who you are) and emotional state (how you feel) (Badcock & Chhabra, 2013). The human voice, therefore, is a rich source of information that extends beyond mere verbal communication. In other words, it encompasses prosodic features that influence perceptions of the speaker’s power, dominance, and emotional state. In the context of leadership, these prosodic features are instrumental.

Over the past century, scholarly attention and numerous theoretical frameworks have been directed toward the exploration of leadership. Among these, transformational leadership has emerged as a particularly noteworthy leadership paradigm in recent decades (Judge & Piccolo, 2004). Transformational leadership embraces an idealistic, positive view of the world, shares strong aspirations, centers supporters' interest on an abstract, long-term vision, encourages progress, and facilitates new ways of working (Yukl, 2010). Leaders who effectively use prosodic features such as pitch, tone, and rhythm can convey confidence, authority, and emotional resonance, thereby enhancing their followers' perceptions of their leadership abilities. Moreover, the significant impact of human speech features on listeners' perceptions suggests that these features play a crucial role in how leadership is perceived. For instance, leaders often exhibit distinct speech patterns that differentiate them from followers, with those in power tending to speak differently (Ko et al., 2015). While evolutionary researchers have emphasized the influence of a low-pitched voice and dynamic vocal spectra between males and females (Hodges-Simeon et al., 2010), it is clear that human speech varies across many dimensions, each contributing to the perception of leadership. Briefly, speech prosody is intricately linked to leadership. The auditory characteristics of speech, particularly non-verbal cues, significantly shape how leaders are perceived by their followers. By harnessing these prosodic features, leaders can enhance their transformational leadership qualities, fostering stronger connections and more effective communication with their followers.

Literature Review

The Role of Fundamental Frequency

Multiple sources have suggested that prosodic features are informative cues to hierarchical organizations and societies (Dunbar & Burgoon, 2005). Besides, these features are critical to reflecting and detecting a hierarchy (Ko et al., 2015). For example, fundamental frequency, also called the voice pitch (F_0), was tested in many studies regarding its influence on perceptions. Klofstad (2016) has found that in an empirical analysis of the 2012 U.S. House of Representatives elections, male candidates with deeper voices received a larger share of votes when facing male opponents. The experiments conducted by Laustsen et al. (2015) revealed that politicians in the United States with deeper voices are voted for more by conservative Republicans than liberal Democrats. Therefore, scholars (e.g., Anderson & Klofstad, 2012; Klofstad, 2016; Tigue et al., 2012) have used fundamental frequency as a perceptual indicator to examine perceived leadership, power as well as dominance. Studies have demonstrated that men with lower-pitched voices were perceived as more attractive (Jones et al., 2010) than those with higher-pitched voices. In contrast, women with higher-pitched voices are perceived as more attractive (Feinberg et al., 2008), whereas females with lower-pitched voices are perceived as more dominant (Borkowska & Pawlowski, 2011).

Research Gap: Speech Prosody and Transformational Leadership

Nevertheless, to the best of the current authors' knowledge, limited literature discusses the relationships between speech prosody and transformational leadership. Based on the Full-Range Leadership Model (Bass, 1985), transformational leadership is identified by four aspects: (1) idealized influence, which is also termed charisma, is the capacity to exercise control by posing as a role model, arousing the pride of followers, and displaying high ethical and moral expectations of behavior; (2) inspirational motivation, which describes the relationship between the leader and followers. By offering meaning and challenge, the leaders empower and encourage followers around them (Bass & Riggio, 2006); (3) intellectual stimulation, which the leader provides, inspires followers to question old assumptions, traditions, and beliefs and to rethink ideas in new ways. The leader challenges followers to be innovative and creative (Bass & Riggio, 2006) to develop their abilities, and (4) individualized consideration, which ensures that by serving as a coach or mentor, the leader pays close attention to the needs of a particular follower for achievement and development (Bass & Riggio, 2006).

Aharonson et al. (2023) have designed a scenario within an educational setting to test the perceived charisma of lecturers, especially focusing on idealized influence. A total of 900 students are instructed to listen to English audio files comprising lectures delivered by 100 male and 100 female speakers. Employing a Random Forest classifier, it has indicated that voicing probability, fundamental frequency, and speech intensity were prominent features associated with charisma. However, the significance of speech intensity in lectures was comparatively diminished. Conversely, the voicing probability- the features linked to rhythm-contributed to creating an engaging pattern in students' perception of lecturers' charisma. Similarly, Feese et al. (2011) conducted a study involving 165 subjects participating in small group meetings to investigate perceptions of individualized consideration, a subdimension of transformational leadership. In the study, all prosodic features were excluded because of an unequal distribution of male and female leaders. Using automatically extracted features and logistic regression, it was found that speech intensity, speaking time, and short utterances were the predictive vocal cues on individualized consideration.

Taken together, the fundamental frequency is a crucial feature in many previous works. While Armstrong et al. (2019) have argued that most relationships are modest and are significantly weaker than the effects on perceptions in laboratory studies. Hence, an interesting question is raised: Is voice pitch a more precise cue? Based on the previous studies (Aharonson et al., 2023; Feese et al., 2011) focused on the subscales of transformational leadership, speech intensity is also an effective feature. The intensity (also known as speech energy, speech power, speech loudness, or sound pressure) of a voice represents the flow of energy through a unit area (Howard & Angus, 2009) and can be physically detected through the pressure of sounds or a subjective level of loudness (Rong & Chen, 2009). Speech intensity grows with the physically measurable sound amplitude, but there is no linear connection, similar to the relationship between fundamental frequency and pitch perception (Neppert, 1999). Thus, there are several complex and controversial psychoacoustic relationships between the volume sensation of a sound and its frequency and duration. According to Vieregge (1996), loudness perception depends on the physically measurable intensity and the subjective overall impression of the speaking voice, the duration of the speech event, the vowel quantities, and the auditory-visually perceptible speaking effort. In addition to determining the average intensity, intensity changes can be scaled in terms of frequency, size, and shape. Apart from fundamental frequency and intensity, voicing probability plays an important role in the study. A voicing probability presents a speech signal's percentage of unvoiced and voiced energy (Gobl & Chasaide, 2003). In every time frame, voicing probability assigns a value that denotes the probability that speech occurs in that frame (Papadopoulos et al., 2016). Higher values of voicing probability indicate speech presence, while lower indicate speech absence (Papadopoulos et al., 2014).

Despite the robust findings in prior works, there is a gap in the existing literature concerning the relationship between speech prosody and transformational leadership. Transformational leadership, as elaborated by Bass and Riggio (2006), consists of four dimensions: idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration. While previous studies have explored the aspects of idealized influence (Aharonson et al., 2023) and individualized consideration (Feese et al., 2011), they have not adequately addressed the full dimensions of transformational leadership, particularly within business settings. Existing works show the impact of the scenarios of lectures and small group meetings, but the field lacks a comprehensive understanding of other contexts. This leads to the inability to provide a comprehensive analysis of how prosody influences transformational leadership perception.

Building upon the foundational work of previous studies (Aharonson et al., 2023; Feese et al., 2011), which have identified specific prosodic features influencing leadership perceptions, our study seeks to provide a more holistic understanding by expanding the scope of analysis to include all four dimensions of transformational leadership. Furthermore, our study investigated the complex psychoacoustic relationships between prosodic features and their combined effect on leadership perception. This nuanced approach is essential as the perception of prosodic features is influenced by multiple factors, such as the physical properties of the voice and the subjective impressions of the listeners. In summary, the existing literature has laid the groundwork by identifying key prosodic features that influence perceptions of leadership. Our study aims to fill this gap by providing a comprehensive analysis of how speech prosody affects all dimensions of transformational leadership, thereby contributing insights to both leadership theory and practical application.

Methodology and Speech Corpus

Participants as Speakers

The research participants served as speakers and were recruited from two universities and a telecommunications company in Germany. Some participants received course credits in exchange for their involvement, while others were provided with incentives as compensation for their participation. Prospective participants underwent screening for factors potentially affecting the quality of their natural voice and were excluded from the study if any such factors were present (Hughes et al., 2014). Prior to the recording sessions, participants provided informed consent and completed a demographic questionnaire capturing information on gender, age, education level, native language, and place of residence. The final sample comprised 122 males, each of whom contributed two speech recording samples utilized as stimuli for the study. The participants had a mean age of 31.48 years (range =20-63 years old, *SD* = 10.48). Regarding age, three participants were less than 20 years old, 49 participants were between 21 and 25 years old, 24 participants were between 26 and 30 years old, 11 participants were between 31 and 35 years old, seven participants were between 36 and 40 years old, another seven participants were between 41 and 45 years old, 15 participants were between 46 and 50 years old, two participants were between 51 and 55 years old, and four participants were over 56 years old. Of these participants, 112 were German, three were Turkish, one was Italian, two were Spanish, one was Russian, two were Polish, and one was French. In addition to the native German speakers, the other participating members with a bilingual background spoke fluent German at the mother tongue level. In terms of educational background, the participants included ten individuals who had completed German secondary school, 47 who had attended German grammar school, and 65 who had pursued higher education at colleges or universities. All 122 participants resided in Germany.

This research focused exclusively on male speakers for two reasons. Firstly, both men and women tend to modulate their voice pitch in various social contexts in relation to factors such as dominance and authority (Sorokowski et al., 2019). Secondly, just 3% of CEO positions were held by females in a study of 6500 venture capital investments (Brush et al., 2014). Therefore, the gender bias in sampling was deemed irrelevant for this research.

Each speaker was presented with a vignette in German, describing an event at a fictional corporation, followed by a voice recording of a CEO. The vignette experiment, also known as a scenario experiment, presents hypothetical situations to research subjects for their responses. Vignette studies are versatile and can be conducted in various settings, including laboratory, field, and online environments (Rietzschel et al., 2017). The use of a fictional corporation was sanctioned to mitigate potential confounding effects arising from participants' prior experiences and perceptions of credibility (Laufer & Jung, 2010). The vignette detailed a scenario wherein the participant assumed the role of the CEO of a paper company overseeing an important project, "Paper for People," and deciding to deliver a speech to inaugurate the project. The speech vignette was adapted from a previous study (Kirkpatrick & Locke, 1996) for the current research. All speakers delivered the same speech content. This approach was adopted to control and minimize the effects of content variability on the results. This standardization allows us to attribute differences in leadership perception more confidently to variations in prosodic features rather than differences in speech content.

The average duration of each speech recording was about 3.5 minutes. The recordings took place individually in a quiet room within a university laboratory. The experimenter was absent from the room during the sessions to mitigate potential Hawthorne effects. Participants were instructed to provide a sample of their normal voice to minimize carry-over effects (Hughes et al., 2014). Audio-visual recordings were captured using a full HD resolution camcorder (Sony HDR-CX 240) mounted on a tripod, with participants wearing a clip-on microphone (Speedlink SL-8691-SBK-01) to enhance voice quality. To avoid additional compression artifacts, audio files extracted from videos were encoded in the uncompressed Waveform Audio Format (.wav) using GoldWave software (Version 5.67, GoldWave, Inc., 2012) with a sampling rate of 16kHz.

Independent Evaluators

Another group of 122 participants, 55 males and 67 females, served as evaluators for the speech samples. The mean age of these evaluators was 30.1 years (range = 19-74 years old, *SD* = 10.9). They were recruited from two German universities and encompassed diverse professional backgrounds, including students and staff. This selection ensured a wide range of perspectives and expertise that led to the robustness of the evaluation process. The inclusion of individuals from various disciplines and roles aimed to reduce potential biases and enhance the generalizability of the findings. None of the evaluators reported hearing impairments that could influence their assessment of the speech samples. Additionally, evaluators were instructed not to proceed with the evaluation if they recognized any speakers. Before beginning the evaluation process, independent evaluators were required to read and sign an informed consent form. Subsequently, they completed a demographic questionnaire, providing information on their gender and age. The evaluation task was conducted on individual computers equipped with headsets (Speedlink SL-8727-BK-01). The whole process took place in a quiet room in the laboratory of a university.

Evaluation Procedure and Instrument: The Multifactor Leadership Questionnaire (MLQ)

The rating procedure was meticulously structured. The timeline of the estimation task included several phases: Initially, 10 minutes were allocated for welcoming the evaluators, introducing the experiment, and completing the consent form. This was followed by a 5-minute technical check of the computer and headset to ensure all equipment was functioning correctly. Evaluators were then given five minutes to understand the rating task through a detailed description. After this, a 3-minute demographic questionnaire was completed to gather essential participant information. A final technical check was conducted over five minutes in the absence of the experimenter to ensure everything was set up. Evaluators then spent five minutes on the speech rating task, followed by ten minutes on the Multifactor Leadership Questionnaire (MLQ) containing the measurable items for transformational leadership perception. After that, the evaluators had a short 5-minute break. Finally, in the presence of the experimenter, the last five minutes were used to take care of technical equipment, express appreciation to the participants, and dismiss them.

The German MLQ scale, which involved translation, modification, and validation processes (see Felfe & Goihl, 2002), was chosen to fit the German evaluators in this study in order to ensure no language barrier. Compared to the original English version (Bass & Avolio, 1995), the German translation demonstrates strong psychometric properties and similar correlations with multiple result parameters (Felfe, 2006). The instrument of the MLO was measured using a 5-point Likert scale, with 1 (strongly disagree) and 5 (strongly agree) for 20 items. Overall, the scales give high reliability (Nunnally, 1978) and present a promising internal consistency measure since the value of total Cronbach's α was greater than .60 (Downing, 2004), as shown in Table 1.

Table 1

Overview of the reliability of the German MLQ (N=122)

Subdimensions	Items	*M*	*SD*	Cronbach's α
Idealized influence attributed	4	3.63	.84	.86
Idealized influence behavior	4	3.83	.77	.79
Individualized consideration	4	3.62	.84	.89
Intellectual stimulation	4	3.67	.81	.85
Inspirational motivation	4	4.05	.93	.92

Analysis of Prosodic Features

Feature Extraction

The prosodic features were computed using the OpenSMILE. OpenSMILE (Version 2.1.0; Eyben et al., 2013) is an open-source toolkit for feature extraction and can compute general-purpose acoustic and prosodic features. It was originally presented in 2009 as an affect and emotion recognition toolkit named OpenEAR (Eyben et al., 2009). One year later, the first upgraded independent release of OpenSMILE (version 1.0.0) was produced (Eyben et al., 2010), and it gradually became one of the most commonly used software programs.

The analysis of prosodic features was conducted using a systematic approach. The speech signal was segmented into frames of 25 milliseconds in length, with a time step of 10 milliseconds. This segmentation process ensured that the temporal resolution of the speech signal was sufficiently fine-grained to capture the dynamic variations in prosody. Following segmentation, a 39-dimensional feature vector was computed for each frame. This feature vector encompassed a comprehensive set of prosodic characteristics essential for the study, including measurements related to the fundamental frequency (voice pitch; F_0), speech intensity, and voicing probability, which the current authors have pointed out as the research gap. In this study, the feature sets and related descriptions were presented in Table 2. To facilitate this process, OpenSMILE (Version 2.1.0) was employed. The default configuration of OpenSMILE was utilized, incorporating well-established algorithms for the extraction of prosodic features.

Table 2

Feature Sets Used in Prosodic Analysis

Feature sets	Prosodic Features and Descriptions	# of Features
Fundamental frequency (voice pitch, F_0)	• F0: Fundamental frequency • F0env: Fundamental frequency of amplitudes	234
Intensity	• pcm_LOGenergy: Signal energy; Mean squared amplitude within a time segment (volume)	117
Voicing probability	• voiceProb: Voicing probability; The proportion of words and pauses which refer to the probability	117

In order to provide a better understanding of the analysis process, authenticated samples of the prosodic feature extraction were provided. Table 3 presented examples of prosodic features with descriptions (c.f. Li et al., 2021; Pfister & Robinson, 2011) of each feature set for three speech samples. These samples were selected to demonstrate the range and variability of the features across different speakers.

Table 3

Example Measurements of Extracted Prosodic Features from Speech Samples

Feature sets	Prosodic features	Descriptions	Speaker 1	Speaker 2	Speaker 3
Fundamental frequency (voice pitch, F_0)	F0_sma_range	The range (max-min) of F_0	470.60	193.2	317.8
	F0_sma_amean	The arithmetic mean of F_0	5.2	0.3	11.1
	F0_sma_stddev	The standard deviation of F_0	35.0	5.9	36.4
	F0env_sma_range	The range (max-min) of F_0 amplitudes	429.2	185.4	302.9
	F0env_sma_amean	The arithmetic mean of F_0 amplitudes	312.6	106.5	155.7
	F0env_sma_stddev	The standard deviation of F_0 amplitudes	68.6	66.2	22.6
Intensity	pcm_LOGenergy_sma_maxPos	The absolute temporal offset of the maximum value of signal energy	9642.0	22232.0	18530.0
	pcm_LOGenergy_sma_absmea	The arithmetic mean of absolute values of signal energy	11.4	22.1	14.0
	pcm_LOGenergy_sma_kurtosis	The kurtosis (4th order central moment) of signal energy	26.9	1.7	7.2
Voicing probability	voiceProb_sma_skewness	The skewness (3rd order central moment) of voicing probability	0.4	0.2	0.1
	voiceProb_sma_numPeaks	The number of peaks of voicing probability	669.0	537.0	758.0
	voiceProb_sma_qregerrA	Linear error between contour and quadratic regression line of voicing probability	1783.0	2317.5	2793.1

Machine Learning Approach

The Waikato Environment for Knowledge Analysis (WEKA) originated from the University of Waikato in New Zealand as a data-mining program encompassing a suite of machine-learning algorithms (Witten et al., 2011) was implemented in this work. It offers comprehensive support for various tasks including data preprocessing, classification, regression, clustering, etc. The present study employs version 3.8.0 of the software.

In recent years, the algorithms of speaker adaption have been proposed and widely applied to speech recognition. The main issue of the adaptation algorithm is that it has to modify many features with only a small amount of adaptation data. The speaker adaptation technology tries to obtain performance close to the speaker-dependent model but has only small amounts of specific data, often based on the initial speaker-independent model. The most commonly used speaker adaptation technology is speaker normalization (Giuliani et al., 2004). This technology normalizes input speech to match the acoustic model to eliminate the different speech features of speakers for training a speaker-independent model. The test data experiences normalization in the training process as well. Thus, both training data and test data are normalized using speaker normalization technology to eliminate the mismatch situation. In general, the speaker normalization has three models: global, iterative, and background. Normalization of global speakers measures normalizing statistics on all data per speaker. In contrast, the iterative speaker normalization method normalizes data by approximating the baseline statistics and background speaker normalization. Iterative speaker normalization automatically calculates the baseline of neutral samples using a recursive algorithm. Unlike the interactive speaker normalization, the background speaker normalization uses a collection of “control” data, often of a distinct lexical content (see Bone et al., 2014).

Therefore, among the three, the background normalization model is most suitable for live applications in this study. In the field of speech processing, speaker normalization is commonly used to improve performance; however, not all speaker normalization methods are practical for real-world applications (Bone et al., 2014). The distribution of basic features shared by all speakers is modeled by the background technique. For individual-level models, this can be a reasonable initial distribution and improves the precision while training data is reduced (Yin et al., 2008). For instance, it may not be easy to have representative data of a participant in the detection of speaker state. Nevertheless, it is possible to access speakers’ neutral or baseline speech in similar recording conditions. In this research, the dataset is split equally, 50%:50%, as training and test datasets. The first 50% of speakers (participant number 1-61) are targeted as a training dataset, and the remaining 50% of speakers (participant number 62-122) are used as a test dataset.

Machine Learning Algorithm

In 1997, the optimization of Support Vector Machines (SVM) was achieved through the decomposition of a complex Quadratic Programming (QP) problem into smaller subproblems (Osuna et al., 1997). Quadratic programming is a foundational technique in mathematical optimization employed to minimize or maximize quadratic functions under linear constraints. Consequently, QP emerges as a common challenge encountered in training SVM models. Subsequent research indicated that the computation of these subproblems is significantly more efficient than solving the entire QP. Thus, the Sequential Minimal Optimization (SMO) algorithm was introduced in 1998 as a promising approach to decompose the problem (Platt, 1998). A distinguishing feature of SMO algorithms is their capacity to address the QP problem by tackling subproblems of size two, each of which possesses an analytical solution (Flake & Lawrence, 2002). Notably, SMO represents the first instance where SVM optimization could be accomplished without reliance on a QP solver. Compared to alternative methods (Joachims, 1999), SMO stands out as the sole online SVM optimizer that directly exploits the quadratic structure of the objective function and effectively leverages two-case size analytical solutions simultaneously (Flake & Lawrence, 2002). Hence, the SMO method among the SVM regression was selected as the algorithm for this research. SMO is fast in overcoming broad QP problems. SMO divides QP problems into a set of simpler QP problems and then analytically solves these small components (Ziari et al., 2016).

Evaluation Metric for Regression Task

In regression analysis, the Root Mean Square Error (RMSE) is a quadratic scoring rule that calculates the error's average magnitude. It is the square root of the average squared differences between actual and prediction measurements. In addition to RMSE, the mean absolute error (MAE) is another principal and widely used measure. MAE measures the generalization error in several predictions. It is regarded as the average of the absolute differences between actual and predicted measurements over the test sample, where all individual differences are equal in weight. Both RMSE and MAE are frequently employed in model evaluation (Chai & Draxler, 2014; Yildiz et al., 2017); a smaller value of them indicates a better performance (Ziari et al., 2016). The previous study discussed the advantages of using MAE over RMSE (Willmott & Matsuura, 2005), while others preferred using RMSE over MAE (Chai & Draxler, 2014). To assess the performance, the present author decided to administer both metrics and the correlation coefficient (R^2) as the evaluation indicators in this study.

Results

Inspirational Motivation

Inspirational motivation assesses the leader's ability in a visual language that can be shared with subordinates to point forward and demonstrate a view of the future. The SMOreg algorithm of WEKA is implemented, with a complexity of 0.1 for SVM. R^2 is a convincing baseline for regression of the best single feature (Hantke et al., 2016). Table 4 illustrates the results for inspirational motivation. Fundamental frequency is a promising feature with R^2 = .31. This implies that fundamental frequency emerges as a pivotal predictor for the subset of inspirational motivation in leadership. This suggests that the voice pitch plays a critical role in conveying future visions and motivating followers. On the other hand, it is noteworthy that voicing probability demonstrates a significant impact, with an R^2 value of .20.

Table 4

Results of SMO Regression on Inspirational Motivation

Feature sets	R^2	MAE	RMSE
Fundamental frequency	.31	.71	.97
Intensity	.16	.75	.95
Voicing probability	.20	.74	.99

Idealized Influence Attributed

In general, idealized influence is the capacity to exercise control by posing as a role model, arousing the pride of followers, and displaying high ethical and moral expectations of behavior. As idealized influence reflects the leader's behavior aspects and attributional components for followers, it can be further separated into two subdimensions: idealized influence attributed and idealized influence behavior. Idealized influence attributed measures a leader's ability to communicate effectively and consistently, coupled with their vocal characteristics, significantly influences how they are perceived as role models by their followers. Table 5 shows the results of SMO regression on the idealized influence attributed. With an exceptionally high R^2 = .45, the regression result of fundamental frequency is likely the best, and intensity is found to be a good feature with R^2 = .36. These findings underscore the importance of both the intensity and the voice pitch in shaping perceptions of idealized influence among followers. Moreover, the voicing probability with R^2 = .31 also plays a role.

Table 5

Results of SMO Regression on Idealized Influence Attributed

Feature sets	R^2	MAE	RMSE
Fundamental frequency	.45	.63	.79
Intensity	.36	.65	.84
Voicing probability	.31	.65	.85

Idealized Influence Behavior

Different from idealized influence attributed, the concept of idealized influence behavior examines the followers' observations of the leader's behavior. Table 6 gives an overview of the SMO regression on idealized influence behavior. Voice pitch and speech intensity are the main reliable predictors with R^2 = .34 and .29. Additionally, while voicing probability may not be as strong a predictor as voice pitch and intensity, its role in shaping followers' perceptions of idealized influence behavior should not be overlooked.

Table 6

Results of SMO Regression on Idealized Influence Behavior

Feature sets	R^2	MAE	RMSE
Fundamental frequency	.34	.59	.78
Intensity	.29	.66	.87
Voicing probability	.23	.60	.82

Intellectual Stimulation

Intellectual stimulation measures the degree to which the leader strives to find fresh solutions, proposes fresh thoughts, and motivates subordinates to challenge or question conclusions and reframe issues. Table 7 provides an overview of regression results for intellectual stimulation. The voice pitch is highlighted with R^2 = .33. In addition, the speech intensity and voicing probability of the predictive features sets are R^2 = .29. These findings underscore the importance of prosodic features in facilitating intellectual stimulation within leadership. Leaders who effectively utilize voice pitch, intensity, and voicing probability are better positioned to foster an environment conducive to innovation and creative problem-solving among their followers. By leveraging these prosodic features, leaders can encourage critical thinking, inspire novel ideas, and drive organizational growth and success.

Table 7

Results of SMO Regression on Intellectual Stimulation

Feature sets	R^2	MAE	RMSE
Fundamental frequency	.33	.59	.79
Intensity	.29	.60	.80
Voicing probability	.29	.59	.80

Individualized Consideration

Individualized consideration reflects whether a leader considers, listens to, trains, and supports subordinates as individuals. The leader is acting as a mentor or coach. Table 8 presents the SMO regression results on individualized consideration. For the specific feature set, the best sign is the speech intensity with $R^2 = .35$, which plays the most important role. By contrast, the most commonly used feature, fundamental frequency results in R^2 values of .22 as well as voicing probability.

Table 8

Results of SMO Regression on Individualized Consideration

Feature sets	R^2	MAE	RMSE
Fundamental frequency	.22	.66	.87
Intensity	.35	.61	.86
Voicing probability	.22	.62	.87

Discussion

Main Conclusion

This study conducted an experiment consisting of 122 speakers and 122 evaluators to examine the transformational leadership perception by prosodic features with SMO regression. Generally speaking, the prosodic features selected in this study, such as fundamental frequency, intensity, and voicing probability, are promising indicators to examine the perceived transformational leadership.

In accordance with Aharonson et al. (2023), the voicing probability, speech intensity, and voice pitch served as vocal cues for charisma, particularly in the domain of idealized influence. However, the importance of the loudness appears to decrease, suggesting that it has little impact in lecture scenarios. In contrast, our research shows the impact of speech intensity in both subscales of the idealized influence, displaying its relationship with dominant or powerful behaviors in the business. Furthermore, our result is consistent with Feese et al. (2011), that speech intensity has been considered an important predictor associated with individual consideration. Individually considerate leaders tend to show greater variability in speech intensity, speak less, and have shorter utterances. These vocal cues of typical individually considerate leaders are related to effective listening (Bass & Reggio, 2006).

Intellectual stimulation requires activities that inspire followers to rethink challenges, create fresh solutions, and experience old circumstances in new ways (Bass et al., 2003). Prior evidence showed a positive relationship between extroversion and intellectual stimulation (Deinert et al., 2015), while extroverted individuals tend to embrace change (Bono & Judge, 2004). According to past research, extroversion perception is influenced by speech pauses (Carney et al., 2007) and speech intensity (Pianesi et al., 2008). Consequently, our findings have demonstrated that voice pitch, intensity, and voicing probability are related to the capacity to foster intellectual stimulation within leadership positions.

Finally, inspiring motivation is the cornerstone of transformational leadership, which emphasizes the importance of developing visions and establishing objectives (Stump et al., 2016). Emotions play an important role in generating motivational impact. By projecting a positive emotion, the speaker will place the listener in a positive state, which in turn affects the desire of the recipient to be influenced positively (Abele, 1999). Generally speaking, the acoustic features of conveying emotions include amplitude, timing, and fundamental frequency (Schirmer & Kotz, 2006). Following this concept, it is postulated that leaders with inspirational motivation should act in a similar way. Some studies have discussed that the lower voice pitch (Weinstein et al., 2018) and softer voice (Hall et al., 2005) were predominant means to influence people. Likewise, the authors found that there is a relationship between voice pitch, voicing probability, and inspirational motivation. Although the intensity is relatively weak, it cannot be overlooked.

In conclusion, our study not only addresses research gaps but also provides a comprehensive examination of the influence of prosodic features on transformational leadership measured by the Multifactor Leadership Questionnaire (MLQ). To the best of the current authors' knowledge, this study represents the first attempt to employ a machine learning algorithm to analyze prosodic features across the full dimensions of transformational leadership within a business context. Our findings underscore the significance of speech prosody in shaping perceptions of leadership and highlight the predictive prosodic features, emphasizing their relevance in understanding leadership dynamics, a domain that has traditionally received more attention in the realm of political leadership studies.

Practice Implications

According to existing literature, charismatic leaders often excel in conveying their positive or negative emotional states through nonverbal communication (Bono & Ilies, 2006). Previous studies have indicated that dynamic vocal tones are frequently employed as strategic tools (Vrij & Mann, 2005), and charisma can be perceived through the voice (Hogenboom, 2014). Consequently, the outcomes of this study strongly support the proposition that leadership or charisma may be trainable or enhanced, particularly in relation to the speech prosody involved (Niebuhr et al., 2016).

First, bear in mind that the fundamental frequency is the key concept. This finding is consistent with several studies, and it has been shown that the voice pitch is a major paralinguistic signal conveying information relating to the expressiveness of speech. Further, it enhances the voice pitch and the intensity of speech. The findings suggest that it is probably useful to have a high loudness level. On top of that, the advantage of a tense voice is that it allows the individual to speak more loudly and stress for emphasis in a more animated and enthusiastic way. More fluent, confident, and enthusiastic-sounding speakers are perceived to be more charismatic (Novák-Tót et al., 2017). Moreover, avoid using long and/or filled pauses, as the duration of speech concerns the receivers. Furthermore, following the principle of speech variation, speakers should avoid always using the same type of emphatic accentuation during the speech.

The advice suggests achieving charisma in speech, such as talks, speeches, and presentations. It illustrates the importance of diligent preparation and intensive practice aimed at effectively conveying leadership attributes to the receivers. Developing charisma in speech requires a substantial investment of time and effort, with speakers focusing on refining their abilities to articulate leadership qualities convincingly to their listeners across various communication contexts. By metaphorically aligning the presentations with the art of dressage, speakers can elevate their performances in a captivating and multifaceted manner. The term "dressage" is originally from French, meaning "training," with the purpose of instilling obedience in horses to respond to commands from their handlers while emphasizing the horse's impulsion and collection. Analogous to the intricate choreography inherent in dressage, speakers can infuse their speeches with dynamic energy, varied pacing, and engaging narratives. Just as a skilled equestrian guides a horse with grace and precision, speakers can navigate their presentations with fluidity and composure. Furthermore, akin to the harmonious relationship between horse and rider, speakers establish a profound connection with their listeners. Through personal anecdotes, humor, or emotional appeals, they forge deep emotional bonds that amplify the impact of their message.

Limitations and Further Directions

This study has some limitations. First, most participants speakers in this research are university students with limited leadership experience in the workplace. It would help to analyze a more diverse population of speakers to understand if there are relationships between prosodic features and transformational leadership in the general public for a more rigorous test of the outcomes. Second, the findings are restricted to the specific background of the business environment, as the level of testosterone has been synonymous with exhibited dominance (Archer, 2006). Leadership is related to higher testosterone in some occupational settings, such as the military. It was found that people would rather pay more attention to dominant vocal signals during wartime (Tigue et al., 2012). In contrast, the opposite is found during peacetime (van der Meij et al., 2016). Therefore, further research could expand the generalizability to other scenarios. Third, leadership is perceived and expressed differently in other cultural backgrounds (Darioly & Schmid Mast, 2014). The concept of transformational leadership is truly American, even though MLQ has been translated into many languages and has been spread to several countries and cultures. A meta-analysis study has demonstrated that transactional leadership was as or more important for many criteria (Judge & Piccolo, 2004). Thus, future research should consider taking transactional leadership into account. Last, future studies could be set up in an international context to explore whether the prosodic features can be extended across cultures. In this vein, the same research question can be discussed in different languages. For instance, Ding et al. (2018) compared the samples of Chinese and German about their voice preferences for the speaker. It has shown that Chinese listeners prefer more pitch-related metric voices and more tone languages than Germans. On the other hand, German listeners prefer more voice quality than local voice pitch movements compared with Chinese. Summing up, the perceptions of listeners could be influenced by different languages and annotations. It remains open if leadership is related to speaking style that differs across different cultural backgrounds (Scherer & Scherer, 1981). Moreover, the cross-corpora comparison of this data corpus with other language corpora might be of interest in the future.

Declarations

Acknowledgements

Not applicable.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Ethics Approval

Not applicable.

Funding Acknowledgements

Not applicable.

Citation to this article

Hsu, C-C., Krajewski, J., & Felfe, J. (2024). How speech prosody affects leadership perceptions: A machine learning approach. *International Journal of Organizational Leadership, 13*(2), 432-450. https://doi.org/10.33844/ijol.2024.60423

Rights and Permissions

References

Abele, A. (1999). Motivationale Mediatoren von Emotionseinflüssen auf die Leistung: Ein vernachlässigtes Forschungsgebiet [Motivational mediators of emotional influences on performance: A neglected area of research].In M. Jerusalem & R. Pekrun (Eds.), *Emotion, motivation und leistung* (pp. 31-50). Göttingen, Germany: Hogrefe-Verlag.

Aharonson, V., Malachi, A., & Katz-Navon, T. (2023, September). *Affective learning and the charismatic lecturer* [Paper presentation]. In the 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW) (pp. 1-4). IEEE. https://doi.org/10.1109/ACIIW59127.2023.10388125

Anderson, R. C., & Klofstad, C. A. (2012). Preference for leaders with masculine voices holds in the case of feminine leadership roles. *PLoS ONE, 7*(12), e51216. https://doi.org/10.1371/journal.pone.0051216

Archer, J. (2006). Testosterone and human aggression: An evaluation of the challenge hypothesis. *Neuroscience & Biobehavioral Reviews, 30*(3), 319–345. https://doi.org/10.1016/j.neubiorev.2004.12.007

Armstrong, M. M., Lee, A. J., & Feinberg, D. R. (2019). A house of cards: Bias in perception of body size mediates the relationship between voice pitch and perceptions of dominance. *Animal Behaviour, 147*, 43–51. https://doi.org/10.1016/j.anbehav.2018.11.005

Badcock, J. C., & Chhabra, S. (2013). Voices to reckon with: Perceptions of voice identity in clinical and non-clinical voice hearers. *Frontiers in Human Neuroscience, 7*, 114. https://doi.org/10.3389/fnhum.2013.00114

Bass, B. M. (1985). *Leadership and performance beyond expectations*. Free Press.

Bass, B. M., & Avolio, B. J. (1995). *MLQ multifactor leadership questionnaire for research*. Mind Garden.

Bass, B. M., Jung, D. I., Avolio, B. J., & Berson, Y. (2003). Predicting unit performance by assessing transformational and transactional leadership. *Journal of Applied Psychology, 88*, 207–218. https://doi.org/10.1037/0021-9010.88.2.207

Bass, B. M., & Riggio, R. E. (2006). *Transformational leadership* (2nd ed.). Lawrence Erlbaum.

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

Bono, J. E., & Ilies, R. (2006). Charisma, positive emotions and mood contagion. *The Leadership Quarterly*, *17*(4), 317–334. https://doi.org/10.1016/j.leaqua.2006.04.008

Bono, J. E., & Judge, T. A. (2004). Personality and transformational and transactional leadership: A meta-analysis. *Journal of Applied Psychology, 89*, 901–910. https://doi.org/10.1037/0021-9010.89.5.901

Bone, D., Li, M., Black, M. P., & Narayanan, S. S. (2014). Intoxicated speech detection: A fusion framework with speaker-normalized hierarchical functionals and GMM supervectors. *Computer Speech and Language*, *28*(2), 375–391. https://doi.org/10.1016/j.csl.2012.09.004

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

Bradac, J. J., Wiemann, J. M., & Schaefer, K. (1994). The language of control in interpersonal communication. In J. A. Daly & J. M. Wiesmann (Eds.), *Strategic interpersonal communication* (pp. 91-108). Hillsdale, NJ: Lawrence Erlbaum.

Brush, C. G., Greene, P. G., Balachandra, L., & Davis, A. (2014). *Women entrepreneurs 2014: Bridging the gap in venture capital*. The Diana Project. Babson College. Retrieved from https://www.babson.edu/media/babson/site-assets/content-assets/about/academics/centres-and-institutes/blank-institute/global-research/diana-project/diana-project-executivesummary-2014.pdf

Carney, D. R., Colvin, C. R., & Hall, J. A. (2007). A thin slice perspective on the accuracy of first impressions. *Journal of Research in Personality*, *41*(5), 1054–1072. https://doi.org/10.1016/j.jrp.2007.01.004

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature. *Geoscientific Model Development, 7*, 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Darioly, A., & Schmid Mast, M. (2014). The role of nonverbal behavior in leadership: An integrative review. In R. E. Riggio & S. J. Tan (Eds.), *Leader interpersonal and influence skills: The soft skills of leadership* (pp. 73-100). Routledge.

Deinert, A., Homan, A. C., Boer, D., Voelpel, S. C., & Gutermann, D. (2015). Transformational leadership sub-dimensions and their link to leaders' personality and performance. *The Leadership Quarterly*, *26*(6), 1095–1120. https://doi.org/10.1016/j.leaqua.2015.08.001

Ding, H., Hoffmann, R., & Jokisch, O. (2018). Voice preference in German: A cross-linguistic comparison of native and Chinese Listeners. *Proceedings of 28th Conference Elektronische Sprachsignalverarbeitung (ESSV)*.

http://essv2018.de/wpcontent/uploads/2018/03/33_DingHoffmannJokisch_ESSV2018.pdf

Downing, S. M. (2004). Reliability: on the reproducibility of assessment data. *Medical Education, 38*(9), 1006–1012. https://doi.org/10.1111/j.1365-2929.2004.01932.x

Dunbar, N. E., & Burgoon, J. K. (2005). Perceptions of power and interactional dominance in interpersonal relationships. *Journal of Social and Personal Relationships*, *22*(2), 207–233. https://doi.org/10.1177/0265407505050944

Eyben, F., Weninger, F., Gross, F., & Schuller, B. (2013). Recent developments in openSMILE, the Munich open-source multimedia feature extractor. *Proceedings of the 21st ACM International Conference on Multimedia* (pp. 835-838). https://doi.org/10.1145/2502081.2502224

Eyben, F., Wöllmer, M., & Schuller, B. (2009). OpenEAR- Introducing the Munich open-source emotion and affect recognition toolkit. In *3rd International Conference on Affective Computing and Intelligent Interaction and Workshop* (pp. 1–6). https://doi.org/10.1109/ACII.2009.5349350

Eyben, F., Wöllmer, M., & Schuller, B. (2010). openSMILE- The Munich versatile and fast open-source audio feature extractor. *Proceedings of the 18th ACM International Conference on Multimedia* (pp. 1459–1462). https://doi.org/10.1145/1873951.1874246

Feese, S., Muaremi, A., Arnrich, B., Troster, G., Meyer, B., & Jonas, K. (2011, October). *Discriminating individually considerate and authoritarian leaders by speech activity cues* [Paper Presentstion]. In 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing (pp. 1460-1465). IEEE. https://doi.org/10.1109/PASSAT/SocialCom.2011.209

Feinberg, D. R., DeBruine, L. M., Jones, B. C., & Little, A. C. (2008). Correlated preferences for men's facial and vocal masculinity. *Evolution and Human Behavior, 29*(4), 233–241. https://doi.org/10.1016/j.evolhumbehav.2007.12.008

Felfe, J. (2006). Validierung einer deutschen Version des, Multifactor Leadership Questionnaire“ (MLQ 5x Short) von Bass und Avolio (1995). [Validation of a German version of the "Multifactor Leadership Questionnaire" (MLQ Form 5x Short) by Bass and Avolio (1995)]. *Zeitschrift für Arbeits- und Organisationspsychologie*, *50*(2), 61–78. https://doi.org/10.1026/0932-4089.50.2.61

Felfe, J. & Goihl, K. (2002). Deutsche überarbeitete und ergänzte Version des "Multifactor Leadership Questionnaire" (MLQ). [German revised and supplemented version of the Multifactor Leadership Questionnaire]. In A. Glöckner-Rist (Ed.), *ZUMA-Informationssystem. Elektronisches Handbuch sozialwissenschaftlicher Erhebungsinstrumente. Version 5.00*. Mannheim, Deutschland: Zentrum für Umfragen, Methoden und Analysen.

Flake, G. W., & Lawrence, S. (2002). Efficient SVM regression training with SMO. *Machine Learning, 46*, 271–290. https://doi.org/10.1023/A:1012474916001

Giuliani, D., Gerosa, M., & Brugnara, F. (2004). *Speaker normalization through constrained MLLR based transforms* [Paper presentation]. INTERSPEECH 2004 - ICSLP, 8th International Conference on Spoken Language Processing (ICSLP 2004), Jeju Island, South Korea. https://doi.org/10.21437/Interspeech.2004-9

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

Grossmann, T., & Friederici, A. D. (2012). When during development do our brains get tuned to the human voice? *Social Neuroscience*, *7*(4), 369–372. https://doi.org/10.1080/17470919.2011.628758

Hall, J. A., Coats, E. J., & LeBeau, L. S. (2005). Nonverbal behavior and the vertical dimension of social relations: A meta-analysis. *Psychological Bulletin, 131*(6), 898–924. https://doi.org/10.1037/0033-2909.131.6.898

Hantke, S., Weninger, F., Kurle, R., Ringeval, F., Batliner, A., Mousa, A. el-D., & Schuller, B. (2016). I hear you eat and speak: Automatic recognition of eating condition and food type, use-cases, and impact on ASR performance. *PLoS ONE, 11*(5), e0154486. https://doi.org/10.1371/journal.pone.0154486

Hargie, O. (2011). *Skilled interpersonal communication: Research, theory and practice* (5th ed.). Routledge.

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

Hogenboom, M. (2014, October 29). Charismatic speakers ‘manipulate their vocal frequency.’ *BBC News*. https://www.bbc.com/news/science-environment-29800893

Howard, D. M., & Angus, J. A. S. (2009). *Acoustics and psychoacoustics* (4th ed). Elsevier.

Hughes, S. M., Mogilski, J., & Harrison, M. A. (2014). The perception and parameters of intentional voice manipulation. *Journal of Nonverbal Behavior*, *38*(1), 107–127. https://doi.org/10.1007/s10919-013-0163-z

Jacob, H., Kreifelts, B., Brück, C., Erb, M., Hösl, F., & Wildgruber, D. (2012). Cerebral integration of verbal and nonverbal emotional cues: Impact of individual nonverbal dominance. *Neuroimage, 61*(3), 738–747. https://doi.org/10.1016/j.neuroimage.2012.03.085

Joachims, T. (1999). Making large-scale support vector machine learning practical. In B. Schölkopf, C. Burges, & A. Smola (Eds.), *Advances in kernel methods- support vector learning* (pp. 169-184). MIT Press.

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*, 57–62. https://doi.org/10.1016/j.anbehav.2009.10.003

Judge, T. A., & Piccolo, R. F. (2004). Transformational and transactional leadership: A meta-analytic test of their relative validity. *Journal of Applied Psychology*, *89*(5), 755–768. https://doi.org/10.1037/0021-9010.89.5.755

Kirkpatrick, S. A., & Locke, E. A. (1996). Direct and indirect effects of three charismatic leadership components on performance and attitudes. *Journal of Applied Psychology*, *81*, 36–51. https://doi.org/10.1037/0021-9010.81.1.36

Klofstad, C. A. (2016). Candidate voice pitch influences election outcomes. *Political Psychology, 37*(5), 725–738. https://doi.org/10.1111/pops.12280

Ko, S. J., Sadler, M. S., & Galinsky, A. D. (2015). The sound of power: Conveying and detecting hierarchical rank through voice. *Psychological Science, 26*(1), 3–14. https://doi.org/10.1177/0956797614553009

Laufer, D., & Jung, J. M. (2010). Incorporating regulatory focus theory in product recall communications to increase compliance with a product recall. *Public Relations Review, 36*(2), 147–151. https://doi.org/10.1016/j.pubrev.2010.03.004

Laustsen, L., Petersen, M. B., & Klofstad, C. A. (2015). Vote choice, ideology, and social dominance orientation influence preferences for lower pitched voices in political candidates. *Evolutionary Psychology, 13*(3), 1–13. https://doi.org/10.1177/1474704915600576

Li, J., Hasegawa-Johnson, M., & McElwain, N. L. (2021). Analysis of acoustic and voice quality features for the classification of infant and mother vocalizations. *Speech Communication*, *133*, 41–61. https://doi.org/10.1016/j.specom.2021.07.010

Mignault, A., & Chaudhuri, A. (2003). The many faces of a neutral face: Head tilt and perception of dominance and emotion. *Journal of Nonverbal Behavior, 27*(2), 111–132. https://doi.org/10.1023/A:1023914509763

Neppert, J. M. H. (1999). *Elemente einer akustischen phonetik [Elements of acoustic phonetics]*. Helmut Buske Verlag.

Niebuhr, O., Voß, J., & Brem, A. (2016). What makes a charismatic speaker? A computer-based acoustic-prosodic analysis of Steve Jobs tone of voice. *Computers in Human Behavior, 64*, 366–382. https://doi.org/10.1016/j.chb.2016.06.059

Novák-Tót, E., Niebuhr, O., & Chen, A. (2017). A gender bias in the acoustic-melodic features of charismatic speech? *Interspeech 2017*, 2248–2252. https://doi.org/10.21437/Interspeech.2017-1349

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed). McGraw-Hill.

Osuna, E., Freund R., & Girosi, F. (1997). An improved training algorithm for support vector machine. *Proceedings of the 1997 IEEE Workshop on Neural Networks for Signal Processing*, 276–285. https://doi.org/10.1109/NNSP.1997.622408

Papadopoulos, P., Tsiartas, A., Gibson, J., & Narayanan, S. (2014). A supervised signal-to-noise ratio estimation of speech signals. *Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 8237-8241). https://doi.org/10.1109/ICASSP.2014.6855207

Papadopoulos, P., Tsiartas, A., & Narayanan, S. (2016). Long-term SNR estimation of speech signals in known and unknown channel conditions. *IEEE/ACM Transactions on Audio, Speech, and Language Processing, 24*(12), 2495–2506. https://doi.org/10.1109/TASLP.2016.2615240

Pfister, T., & Robinson, P. (2011). Real-time recognition of affective states from nonverbal features of speech and its application for public speaking skill analysis. *IEEE Transactions on Affective Computing*, *2*(2), 66–78. https://doi.org/10.1109/T-AFFC.2011.8

Pianesi, F., Mana, N., Cappelletti, A., Lepri, B., & Zancanaro, M. (2008). Multimodal recognition of personality traits in social interactions. *Proceedings of the 10th International Conference on Multimodal Interfaces* (pp. 53-60). https://doi.org/10.1145/1452392.1452404

Platt, J. (1998). Fast training of support vector machines using sequential minimal optimization. In B. Schölkopf, C. Burges, & A. Smola (Eds.), *Advances in kernel methods—support vector learning*. MIT Press.

Rietzschel, E. F., Wisse B., & Rus, D. (2017). Puppet masters in the lab: Experimental methods in leadership research. In B. Schyns, R. J. Hall, & P. Neves (Eds.), *Handbook of methods in leadership research* (pp. 48-72). Edward Elgar Publishing. https://doi.org/10.4337/9781785367281.00010

Rong, J., Li, G., & Chen, Y-P. P. (2009). Acoustic feature selection for automatic emotion recognition from speech. *Information Processing and Management, 45*(3), 315–328. https://doi.org/10.1016/j.ipm.2008.09.003

Scherer, K. R. (1982). Methods of research on vocal communication: Paradigms and parameters. In K. R. Scherer & P. Ekman (Eds.), *Handbook of methods in nonverbal behavior research* (pp. 136-198). Cambridge University Press.

Scherer, K. R., & Scherer, U. (1981). Speech behavior and personality. In J. K. Darby (Eds.), *Speech evaluation in psychiatry* (pp.115-135). Grune & Stratton.

Schirmer, A., & Kotz, S. (2006). Beyond the right hemisphere: Brain mechanisms mediating vocal emotional processing. *TRENDS in Cognitive Sciences, 10*(1), 24–30. https://doi.org/10.1016/j.tics.2005.11.009

Sorokowski, P., Puts, D., Johnson, J., Żółkiewicz, O., Oleszkiewicz, A., Sorokowska, A., ... & Pisanski, K. (2019). Voice of authority: Professionals lower their vocal frequencies when giving expert advice. *Journal of Nonverbal Behavior*, *43*, 257–269. https://doi.org/10.1007/s10919-019-00307-0

Sporer, S. L., & Schwandt, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology*, *20*(4), 421–446. https://doi.org/10.1002/acp.1190

Stump, M., Zlatkin-Troitschanskaia, O., & Mater, O. (2016). The effects of transformational leadership on teachers' data use. *Journal of Educational Research Online, 8*(3), 80–99. https://doi.org/10.25656/01:12807

Tigue, C. C., Borak, D. J., O'Connor, J. J. M., 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

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

van der Meij, L., Schaveling, J., & van Vugt, M. (2016). Basal testosterone, leadership and dominance: A field study and meta-analysis. *Psychoneuroendocrinology*, *72*, 72–79. https://doi.org/10.1016/j.psyneuen.2016.06.005

Vieregge, W. H. (1996). *Patho-Symbolphonetik: auditive Deskription pathologischer Sprache* [Patho-symbolic phonetics: auditory description of pathological language]. Steiner.

Vrij, A., & Mann, S. (2005). Police use of nonverbal behavior as indicators of deception. In R. E. Riggio & R. S. Feldman (Eds.), *Applications of nonverbal communication*. Lawrence Erlbaum.

Weinstein, N., Zougkou, K., & Paulmann, S. (2018). You 'have' to hear this: Using tone of voice to motivate others. *Journal of Experimental Psychology: Human Perception and Performance, 44*(6), 898–913. https://doi.org/10.1037/xhp0000502

Willmott, C. J. & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the mean square error (RMSE) in assessing average model performance. *Climate Research*, *30*(1), 79–82. https://doi.org/10.3354/cr030079

Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques*. Burlington: Morgan Kaufmann.

Yildiz, B., Bilbao, J. I., & Sproul, A. B. (2017). A review and analysis of regression and machine learning models on commercial building electricity load forecasting. *Renewable and Sustainable Energy Reviews*, *73*, 1104–1122. https://doi.org/10.1016/j.rser.2017.02.023

Yin, B., Chen, F., Ruiz, N., & Ambikairajah, E. (2008). *Speech-based cognitive load monitoring system.* IEEE International Conference on Acoustics, Speech and Signal Processing, 2008 (ICASSP 2008), Las Vegas, USA. https://doi.org/10.1109/ICASSP.2008.4518041

Yukl, G. A. (2010). *Leadership in organizations* (7th ed.). Pearson Hall.

Ziari, H., Maghrebi, M., Ayoubinejad, J., & Waller, S. T. (2016). Prediction of pavement performance: Application of support vector regression with different kernels. *Transportation Research Record: Journal of the Transportation Research Board, 2589*, 135–145. https://doi.org/10.3141/2589-15