An Interdisciplinary Study of A Leaders' Voice Characteristics: Acoustical Analysis and Members' Cognition

SangWoo Hahm[1] and Hyungwoo Park[2*]
[1] Department of Business Administration, Semyung University
65 Semyung-ro, Jecheon, Chungbuk, korea
[e-mail: thehahm@semyung.ac.kr]
[2] Quantum Dot Developemt Team, Samsung Display
1 Samsung-ro Giheung-gu, Yongin-si, Gyeonggi-do, Republic of Korea
[e-mail: hyungwoo.park@samsung.com]
*Corresponding author: Hyungwoo Park

Abstract

The traditional roles of leaders are to influence members and motivate them to achieve shared goals in organizations. However, leaders such as top managers and chief executive officers, in practice, do not always directly meet or influence other company members. In fact, they tend to have the greatest impact on their members through formal speeches, company procedures, and the like. As such, official speech is directly related to the motivation of company employees. In an official speech, not only the contents of the speech, but also the voice characteristics of the speaker have an important influence on listeners, as the different vocal characteristics of a person can have different effects on the listener. Therefore, according to the voice characteristics of a leader, the cognition of the members may change, and, the degree to which the members are influenced and motivated will be different. This study identifies how members may perceive a speech differently according to the different voice characteristics of leaders in formal speeches. Further, different perceptions about voices will influence members' cognition of the leader, for example, in how trustworthy they appear. The study analyzed recorded speeches of leaders, and extracted features of their speaking style through digital speech signal analysis. Then, parameters were extracted and analyzed by the time domain, frequency domain, and spectrogram domain methods. We also analyzed the parameters for use in Natural Language Processing. We investigated which leader's voice characteristics had more influence on members or were more effective on them. A person's voice characteristics can be changed. Therefore, leaders who seek to influence members in formal speeches should have effective voice characteristics to motivate followers.

A preliminary version of this paper was presented at ICONI 2019.

1. Introduction

The majority of definitions of leadership center around it being a process to influence, guide, and facilitate members to achieve common goals in organizations. Moreover,leaders should enhance performance, anticipate the future, and change their organizations positively [1, 2]. The traditional role of leaders is to influence members and motivate them to achieve shared goals in organizations. Leaders can influence followers through a diverse range of traits and behaviours such as reward systems, becoming and acting as a role model, making policy decisions, and conversations [3-5]. However, in practice, leaders such as top managers and chief executive officers do not often directly meet or influence employees. In fact, they have the greatest impact on their members through formal speeches, company procedures, structures and the like. Therefore, official speeches are directly related to the motivation levels of the members and the degree of influence of the leader. In an official speech, not only the content of the speech, but also the voice characteristics of the leader can have an important influence on members, as the different vocal characteristics of a person can have different effects on a listener. Therefore, according to the voice characteristics of the leader, the perceptions of members may change. This means that depending on the voice characteristics the leader uses in a formal speech, the degree to which the members are influenced and motivated will be different. The perception which members have about their leader directly affects their preference and satisfaction levels about the leader. Therefore, not only the voice characteristics of leaders themselves, but also the members' perception of these voice traits, are crucial factors on influence and motivation. Depending on how the members perceive the leader's voice characteristics, they can be more or less motivated [6].

Generally speaking, voices convey people's thoughts and information very easily. The combination of language, speech sounds, analytical composition, frequency distributionof voice, energy change over time domain, and fundamental frequency (pitch) are variable. In addition to transmitting information through language, voices carry additionalvaried information such as emotion, tone, health, favourability, and stress. All voices are different from each other, as each individual's vocal organs are different, and they each have their own unique characteristics, such as vocabulary habits [7-14].

This study identifies different perceptions of members based on the different voice characteristics of leaders in formal speeches. Hence, we investigated which leader's voice characteristics had more influence on the members or were more effective in conveying information to them. A person's voice characteristics can be changed. Therefore, leaders who seek to influence members in formal speeches should possess effective voice characteristics to motivate followers. Members listen to leader's voices in speeches, and they feel, recognize and perceive these voices differently. We set five factors to help understand members' perceptions of leaders' voice characteristics. These cognitive variables can explain how followers feel and think about leaders' voice characteristics [15]. Members will have different perceptions depending on the leader's voice characteristics such as pitch harmonics, formant transitions, and speech rate. Furthermore, we explain how these perceptions affected the motivation of the members. We measure trust in leaders to explain the motivation of the members. Trust in a leader increases members' satisfaction and motivation levels [16]. Therefore, first we clarified the effect of a leader's voice characteristics on the cognitive levels of members in formal speeches. Next, we demonstrated how each cognitive dimension affects trust in the leader. Hence, we showed that certain voice features have a positive effect on member's perception and can increase trust in leaders which is in turn related to worker's motivation, satisfaction and performance.

This research explores the voice characteristics that can improve the trust in a lead e r. For this purpose, interdisciplinary research through voice characteristic analysis and cognition was attempted. These attempts can have a positive effect on the performance of natural language processing (NLP). Natural Language Processing (NLP) is about the overall technique of interactions between computers and human languages in processing. NLP also focuses on more human-oriented interfaces. In NLP, we study meaning, emotion, and emotional cognition as well as the effects of processing on voice, image, and text. In this study, we conducted an interdisciplinary approach on speech signals and cognitive effects on human of NLP. It is posited that a convergence study on voice and cognition can improve performance so that NLP can better understand and respond to human language. For example, NLP may be able to identify how authentic the voice of a particular person is, or it can respond to a person with a more reliable voice via text to speech (TTS) [17-21].

2. Literature Review

2.1 The Linear prediction model of voice generation system

For speech signal processing, various models have been studied. Among them, the linear predictive coding method in the middle of the 20th century is widely used [13]. In this model, the voice signal is largely divided into voiced / unvoiced sound. Then, the analysis results are divided into the generation principle and the propagation process's vocal track parameters. Voice is a time-independent changing signal and has a short periodicity. In addition, it is related to surrounding signals and is called a linear prediction model by utilizing Vocal-track information that can predict the future using the current state [17]. The voice is based on the principle of generation, and the voiced sound uses the impulse train, and its characteristic parameters are the period, magnitude, and rate of change of the impulse train. In the unvoiced sound, information similar to the white sound is used. The vocal track information refers to a set of frequency peaks of transmission information of each pronunciation and is information obtained during the delivery process. As for the resonance information of the vocal track, information about pronunciation appears in the lower third-order and higher, and the speaker's speech characteristics are expressed in the third-order or higher [13, 17-24].
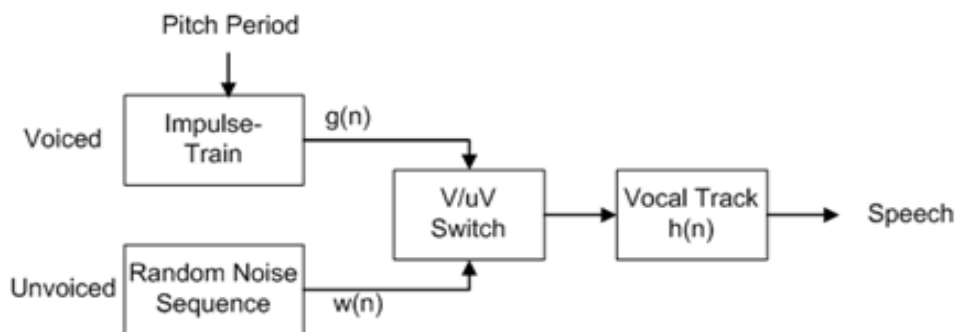

Fig. 1. Speech generation model [17]

The speech model shown in the block diagram in **Fig. 1** has two important parameters. One is the pitch associated with the generation, and the other is the formant associated with vocal track [17]. The pitch makes the vibrations of the air stronger as the air passes through the openings between the vocal folds of the vocal cords. At this time, the observed type of information appears as a pitch [22]. The audio signal changes with time, but in a short period,

this vibration produces quasi-periodic information. The pitch information is expressed by the basic vibration period of the vocal cords and is also converted to fundamental frequency [23, 24].

Second, the important information is the formant. Formant information appears as a prominent bud in the rest of the voice waveform or frequency response curve after excluding the part corresponding to the pitch. This section converts the audio signal in the short section to the cepstrum area, removes the part corresponding to high quefrency, and corresponds to the rest. The flow of the voice signal in **Fig. 1**, the signal suitable for the generation characteristic is obtained through the filter corresponding to the vocal track. At this time, formant information can be used as the transfer function of the filter. Physically, it can be determined that the air trembling generated by the vocal cords passes through the resonance of the vocal organ [25].

2.2 Voice Analysis Methods: Speech Rate

The speech rate is a measurement of the number of phonemes generated within a unit of time, typically the numbers of phonemes per 1 sec. The duration of pauses also determines that in the speech. The rate of speech varies according to many factors, including the speaker's personality, level of knowledge, the region that the speaker is from, the kind of language being used, and the situation the speaker is in. Preliminary studies have revealed that speaking about three to four characters per second helps the listener better hear and comprehend. It has also been proposed that a speaker who speaks quickly is accepted as more persuasive than a slow speaker and gives a more favorable impression in terms of ability and social appeal [13, 17-23].

The methods used to measure speech rates include checking the duration of the phoneme, analysing the LPC (Linear Predictive Coding) and the LSP (Line Spectrum Pair) to distinguish between the duration and the pause, and calculating the speech rate. In previous studies, we measured the speech rate in the context of determining the transmission rate of the speech coder. We used an algorithm to measure the change in formant as a result of the hybrid domain analysis in a short interval, and also measured the change using the LSP [24, 25].

2.3 Literature Review for Cognitive Variables

2.3.1 Cognitive Variables of Members on Formal Speech

Several cognitive variables may influence the way that listeners perceive speech. The first of these is eloquence, which essentially means fluency in speaking. The appropriate level, pronunciation, and tone make a speech more valuable [15, 26]. Pausing appropriately in a speech may also help listeners feel that a speaker has greater command of fluency. A leader's oratorical power may influence members' perception positively in a formal speech. The second factor is amiableness, which is an individual's preference for voice traits. Naturally a favourable impression toward a leader has a positive relationship on a leaders' influence and members' motivation [26]. Further, listeners differ in the degree of comfort, appeal, or negative reaction that they feel depending on the voice characteristics of a leader. The speed, volume, and pitch of a voice that attracts them can also be different [27]. Therefore, the perceived feelings of the listener may differ depending on the voice characteristics of the leader making the speech. Third is authenticity, which refers to owning one's personal experiences captured by the injunction to know oneself. Voice features, such as the speed of speech, influence the authenticity perceived by listeners [15, 28]. For instance, a slow voice tends to sound less authentic and lowers the trustworthiness and ability of the speaker to

persuade. Thus, the voice features of the leader will also affect whether listeners perceive him or her to be authentic or not. Fourthly is clarity, a concept related to ease of understanding. If a leader uses words that are easy to pronounce, or if the pronunciation itself is clear, lively, and expressive, members will more clearly understand the leader's utterances [29, 30]. Therefore, there is a potential difference in how much the members understand according to the degree of clarity or clearness of the leader's voice. Finally, voice quality, speed, rest, and intonation can enhance the effectiveness of communication and persuasion, and also affect the attitude of the audience. Hence, depending on the characteristics of the voice, people are more likely to concentrate or to remember more, which is directly connected to absorption [31, 32]. In short, a leader with positive voice characteristics will be more likely to foster positive attitudes such as commitment and absorption.

2.3.2 Trust in the Leader

Trust has been conceptualized as "a psychological state comprising the intention to accept vulnerability based upon positive expectations of the intentions or behaviour of another" [33]. Moreover, trust is defined as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" [34]. Interpersonal trust is defined as "the expectancy held by an individual or a group that the word, promise, verbal or written statement of another individual or group can be relied upon." [35]. Hence, trust in a leader refers to a members' psychological state of having a willingness to believe, follow and anticipate things from leaders. Moreover, trust in a leader can directly and indirectly increase followers' levels of satisfaction, motivation and performance [26, 29, 36]. There are very few studies on the relationship between the characteristics of a leader's voice and trust in a leader. However, a leader's traits, such as the characteristics of the voice, are related to the satisfaction, motivation, and influence over their members [5]. An individual member can prefer the specific voice characteristics of a leader and be more satisfied with them. Hence, members who are influenced by the leader's voice characteristics will be more satisfied and will trust their leader more. Therefore, we predict that the particular voice characteristics of a leader will affect the members' perception, such as trust in the leader.

2.4 Perspective of Natural Language Processing

NLP is about the overall technique of interactions between computers and human languages in processing. Natural language is commonly used by humans. It is a concept opposed to artificial languages such as programming codes like C and Java. The purpose of NLP is to understand, process, and command human languages, and to translate the results back into a natural language. For NLP, there are various human-computer-interface fields such as voice, video, and text. Today, NLP not only communicates information, but also comprehends human emotions and fuses them with cognitive science to determine contextual information. NLP is divided into rule-based processing and statistical-based processing. Rule-based processing NLP developed in the past due to a lack of computational power. These days, NLP performance has improved rapidly in statistical-based processing due to the development of algorithms such as big data analysis and deep learning, and also GPU processing technics. Rule-based processing shows a high accuracy in the execution of limited commands, but the scalability to general human communication is poor. In recent years, AI has been developed to enable machine learning and high-quality responses based on the development of AI [37-42]. NLP is based on an understanding of human behavior because it requires the use of human language, the comprehension of human emotions, the prediction of human behavior, and

necessary responses. Understanding human behavior is based on the linguistic discipline of learning the characteristics of the language used, and needs to be able to understand not only information, but situations, moods, and feelings, like human-to-human communication. In order to understand this non-linguistic information, previous studies have analyzed certain emotions and used the characteristics expressed as an analysis and utilization factor. However, there is a lack of research on how people actually received this information, so the characteristics of calculated NLP are different from the emotions and feelings of people. In this study, an interdisciplinary study of speech signal processing and cognitive effects was performed. In other words, the core of NLP is human understanding and human-oriented communication [37-41]. Hence, sending (i.e speaking) as well as receiving (i.e listening) is important for communication. Existing research on voices focuses on sending [42]. In one study, common characteristics of actors are extracted to explain what specific feelings (angry, sad, etc.) are being conveyed. However, this is not a characteristic of the voice of the listener. When listening to one person's voice, people can feel or recognize it differently. Thus, we need parameters of the cognition of listeners that can be used in NLP.

3. Experiment and Results

3.1 Participants and measurement

To conduct the analysis, we surveyed 76 employees working as drivers at a transportation company. Three managers of the company gave official speeches on similar topics related to corporate training, such as safety education, vehicle maintenance, and legal guidance. The workers listened to these official speeches and responded to their cognition of the managers' voices. Then they checked the degree of trust towards these managers.
The workers consisted of 52 males (68.4%) and 24 females (31.6%). In terms of age, 12 were in their twenties (15.8%), 28 in their thirties (36.8%), and 36 were older than 40 (47.4%). In terms of academic experience, 38 had completed high school (50%), 28 had bachelor degrees (36.8%), and 10 were masters degree holders (13.2%). For length of employment, 44 had worked less than one year (57.9%), 22 had worked one to three years (28.9%), and 10 had worked longer than three years (13.2%).

All measurements used a 7-point Likert scale (1: strongly disagree, 7: strongly agree). Each cognitive factor had four items. Questionnaires were centred around the variables of eloquence (is impressive, speaks fluently, appeals, speaks well), amiableness (attractiveness, makes me want to listen, speaks cheerfully, speaks comfortably), authenticity (is honest, does not hide, does not deceive, speaks frankly), clarity (clear expression, is not complicated, but simple, speaks directly, speaks easily), and absorption (makes me concentrate, makes me focus, makes me listen well, speech time goes quickly). Trust in the leader was measured with four items, including "I feel a strong loyalty to my leader" and "I would support my leader in almost any emergency" [35, 36].

3.2 Voice Analysis and Results

This paper compared and analysed several parameters of leaders' voices such as confidence, control, and stress level by comparing vocal signals from three conference announcement audio files. The audio file was recorded during the weekly meeting of a town bus company. Each recorded speech data was sampled at 22 kHz, and quantization was 16 bits. For the results of this study, the primary stage voice preprocessing was refined using 'Adobe' 'Audition'. Since then, 'Matlab' was used for data analysis and processing, and provided libraries such as 'Signal

Processing Toolbox' were used to reduce the analysis time. As for the hardware environment, ordinary Intel's i7 processor, 16G RAM. We didn't need parallel computing, so didn't take advantage of GPU acceleration. A 30 ms window in the form of the 'Hamming window' was used for short-term time-domain analysis, and it was continuously analyzed with 30% overlap. In each data, the periodicity was emphasized and extracted for pitch detection. For pitch detection in this study, a frequency correlation method was used in the frequency domain in continuous shart-term data. The pitch of the human voice originally exists in the range of 80 to 500 Hz, and since the voice of a male is processed in particular, the detected pitch other than the 80-350 Hz band is ignored.

Fig. 2 shows a characteristic portion of the vocal spectrum, where the green solid line represents the first leader's voice from the first 10 minute speech, the yellow solid line represents the second leader's voice characteristics and the red line is the analysis of the third leader. It can be seen that the average pitch frequency differed among the three speakers. The slope of F1 from F2 represents the voice affinity.
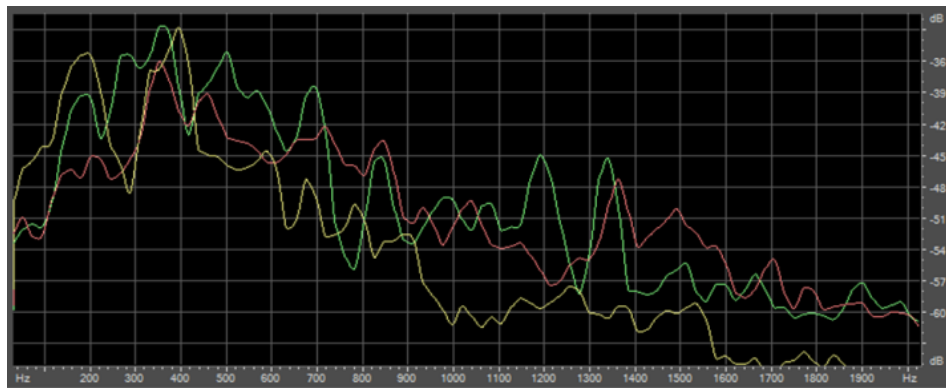

Fig. 2. Speech spectrum analysis

Fig. 3 and **Fig. 4** below show the analysis results of the first leader. **Fig. 3** shows the results of the pitch analysis in the frequency domain, and **Fig. 4** shows the results of the spectrogram analysis. In **Fig. 3**, the fundamental frequency and its harmonic structure demonstrate information that includes the softness of the voices, and this frequency value is in turn determined by the speakers' vocal cords' health and age. Further, when someone recognizes a change in strength of the pitch, they are able to determine information about emotions and feelings. The voice of the first leader had an average pitch tone of 172Hz. This frequency corresponds to the mid-high tone, where the feeling of stability is somewhat lacking. However, this pitch frequency could also convey a smooth first impression. Additionally, if we look at the structure of the pitch, we can see that it is well organized up to 1500 Hz. This factor can positively stimulate the hearing of listeners [13, 14, 18].

To assess good voices objectively, we can parameterize softness and attentiveness. This can be objectified by the harmonic structure and bandwidth of the pitch frequency, and can be evaluated using the degree of tightness of the structure of harmonics and the degree of distinctness between the pitch bands. Finally, an evaluation can be done based on the speed of speaking, which reveals a comfortable speed of listening of 3.5-characters / second. Thus, using the criteria above, the first leader's objective 'good voice' score is 8 out of 10 points. The degree of leadership that a voice rated 8/10 imparts is relatively high – and such a voice can be deemed as being good to listen to [13, 14, 18].
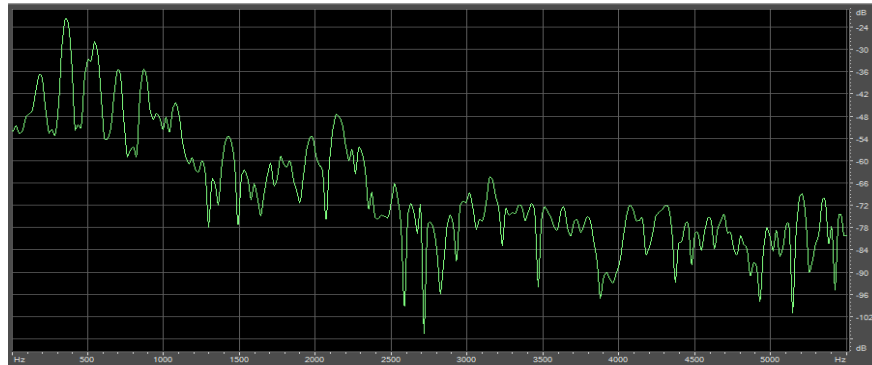
Fig. 3. Spectrum analysis of leader1 voice

Fig. 4 shows the spectrogram analysis of the voice of the first leader. It shows that the pitch structure of the 5th to 6th harmonics appears well from the pitch of the lowermost band, and it can be confirmed that the syllables change clearly. This result means that the speaker has a highly appealing voice. A clear change in formant start and end point of the voice means that a speaker's pronunciation is correct, and this feature becomes a factor helping the listener to more clearly understand what is being said. When we analyse the rate of change of speech according to time, we find that the first speaker speaks at a rate on average of five characters per second, and thus he can be judged as having an adequate speaking speed and energy. Therefore, the pitch structure of the first leader is attentive and has a soft voice characteristic, the spectrogram reveals a clear syllable change, and a clear and appropriate speed and energy, which allows us to infer that the speaker has an appealing voice [13, 14, 18].
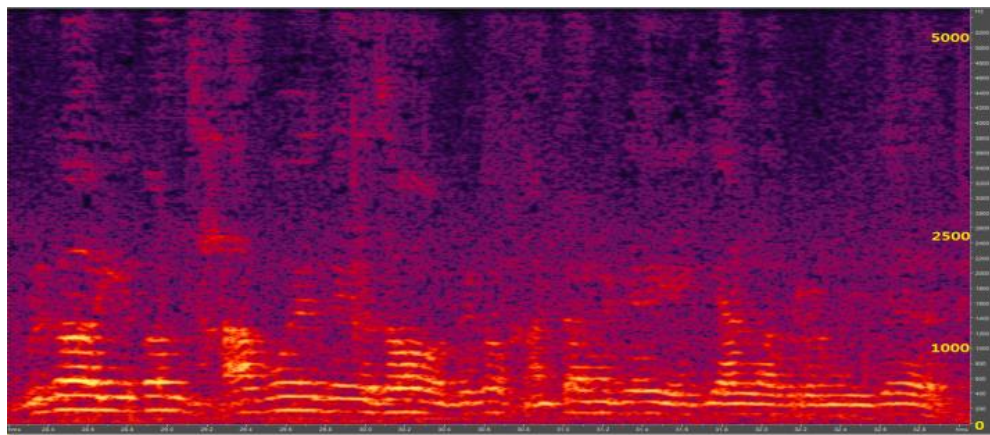

Fig. 4. Spectrogram analysis of leader1 voice

Fig. 5 and **Fig. 6** below show the analysis results for the second leader. **Fig. 5** shows the results of the pitch analysis in the frequency domain, and **Fig. 6** shows the results of the spectrogram analysis. In **Fig. 5**, the fundamental frequency and its harmonic structure can be seen, and this can be evaluated in the same way as the first leader's judgment analysis. The pitch frequency of the second leader is 150 Hz, which corresponds exactly to the mid-bass. However, compared to the first leader, the second leader's voice is less confirmed in that area where the pitch harmonics structure is less, and energy is emphasized only in a specific band. Putting this information together leads us to give Leader 2 a score of 6 points out of 10. What this score indicates is that there is some possibility of the second leader having a good voice,

but as there are few high sections in the harmonic structure, and a high rate of speech after the middle part of the speech, it may be somewhat difficult to understand all of the speech correctly [13, 14, 18].
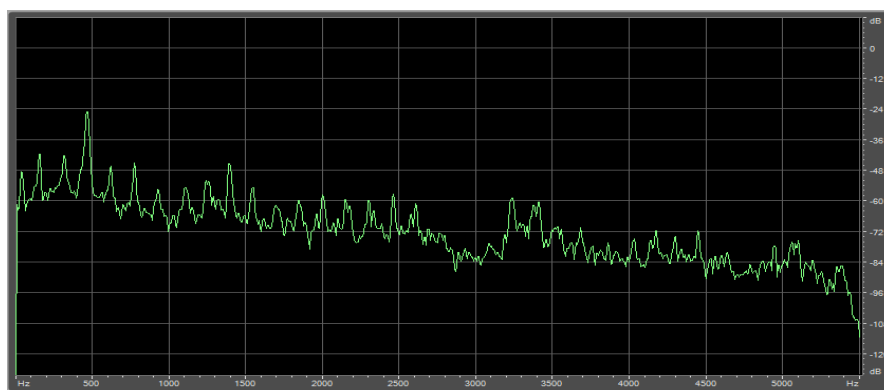

Fig. 5. Spectrum analysis of leader2 voice

Fig. 6 shows the results of an analysis of the second leader's spectrogram. In this analysis, only the pitch of the lowermost band is shown from the pitch to the fifth harmonic, and the harmonics interval is also small. Further, the section indicating syllable change is not clear. These results signify that this speaker has low voice appeal. In addition to this, the distinct disadvantage of a non-prominent formant structure is found in several sections. Taken together, these results indicate that the speaker is hard to follow and understand clearly. The analysis of the rate of change of speech according to time showed an average rate of 7 characters per second, and the voice often became difficult to understand perfectly in the middle of a sentence. This is related to the speed of speech and the need to consistently maintain the average speed of speaking. If syllables cluster or become dispersed, as in the middle of speaker 2s speech, listeners will lose confidence in the voice, and they may disrespect or ignore the speaker [13, 14, 18].
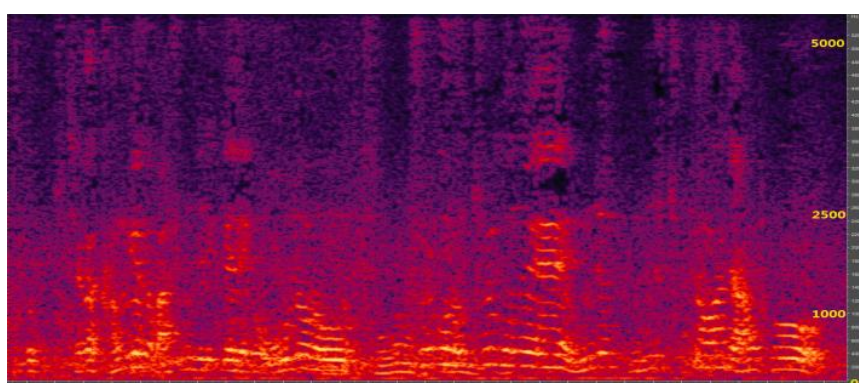

Fig. 6. Spectrogram analysis of leader2 voice

Fig. 7 and **Fig. 8** below show the analysis results for the third leader. **Fig. 7** shows the results of the pitch analysis in the frequency domain, and **Fig. 8** shows the results of the spectrogram analysis. In **Fig. 7**, the fundamental frequency and its harmonic structure can be seen, and it can be evaluated in the same way as the first and second leader's judgment analysis. The fundamental frequency of the third leaders pitch frequency is 107 Hz, which corresponds to a low frequency. And although the harmonic structure of the pitch appears to be of mid-range quality, its structure is uneven, and breathing is not regular in the middle of the sentence. A score for Leader 3 calculated based on the same criteria for evaluating leader 1 or 2 will be 5 points out of 10. There is a possibility of a good voice being heard, but the harmonic structure appears differently at each moment of the utterance. In the middle of the sentence, utterances are repeated and duplicated. This can lead to boredom and discomfort for listeners [13, 14, 18].
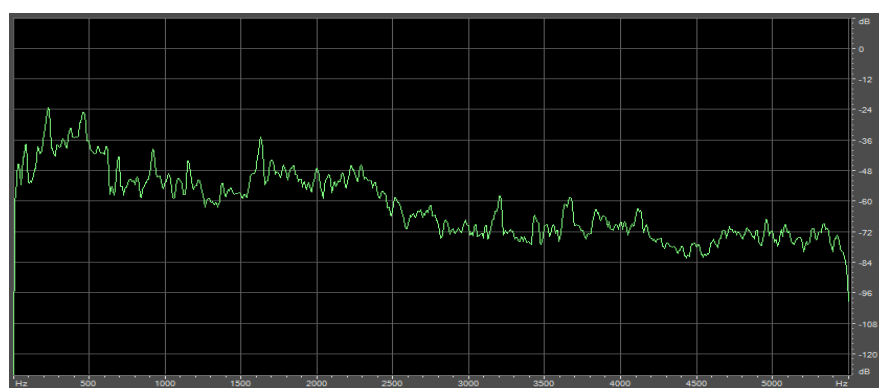

Fig. 7. Spectrum analysis of leader3 voice

Fig. 8 shows the results of a spectrogram analysis of the third leader. It can be seen that the 8th to 9th harmonic structure appears from the pitch of the lowermost band. However, in the section of the syllable change, the formants of the third or higher order were not clear and are blurred. This leads to unclear vocalization, which in turn leads to speaker 3 receiving low scores. Taken together, these results mean that leader 3 is hard to hear and understand clearly, and a person listening to this leader speak could very likely feel stressed or agitated. The analysis of the rate of change of speech according to time shows that rate of speech averages just two characters per second, and that some words in the middle of sentences are repeated. These are hallmarks of a boring and unpleasant speech. As was the case with the second leader, if syllables cluster or loosen in the mid-stages, people lose confidence in the voice they are listening to and tend to dismiss or ignore the speaker [13, 14, 18].
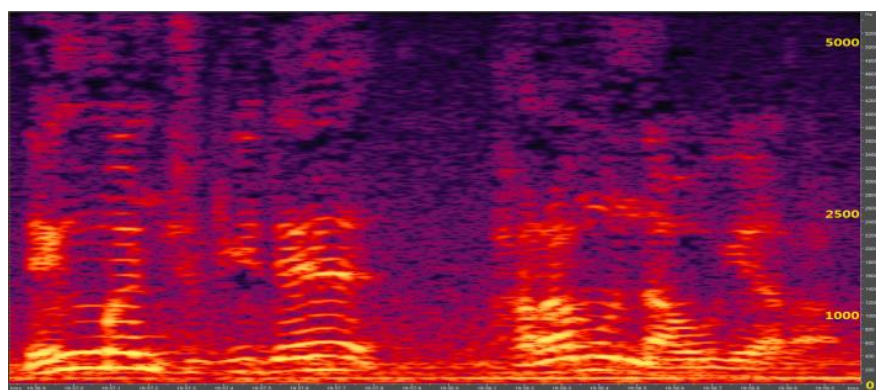

Fig. 8. Spectrogram analysis of leader2 voice

3.3 Statistic Analysis

Table 1 indicates the results of a confirmatory factor analysis (CFA), reliability, and descriptive statistics. As a result of the CFA, all variables were significantly valid, and have enough indexes (AVE >0.5, composite reliability >0.7; and absolute, incremental and parsimonious fit indexes showed sufficient values). Also, all factors' Cronbach's α values were higher than .7, thus reliability has been secured.

Table 1. Result of CFA, reliability and descriptive statistics

	AVE	Composite Reliability	Cronbach's α	Mean	Standard Deviation
eloquence	.644	.808	.921	3.980	1.408
amiableness	.801	.914	.965	3.757	1.501
authenticity	.957	.987	.995	3.917	1.504
clarity	.880	.961	.987	3.816	1.649
absorption	.943	.983	.994	3.504	1.621
trust in leader	.803	.920	.963	3.722	1.333
fit index	absolute	X^2(p)=989.019, df=211, X^2/df =4.687 ($\leq$5.0)			
	incremental	TLI=.912 ($\geq$0.9), CFI=.933 ($\geq$0.9), IFI=.933 ($\geq$0.9)			
	parsimonious	PNFI=.701 ($\geq$0.6), PGFI=.512 ($\geq$0.5)			

Table 2 displays the results of correlation among variables. Most of the factors have positive correlations with the other factors. However, pitch harmonics did not have a significant correlation with absorption, trust in a leader nor the formant transition section; and also trust in leader did not have correlation with formant transition section and speech rate.

Table 2. Results of correlation analysis

	1	2	3	4	5	6	7	8	9
1	-								
2	.000	-							
3	.982***	.189***	-						
4	.278***	.130**	.298***	-					
5	.258***	.170***	.285***	.897***	-				
6	.002	.020	.006	.609***	.565***	-			
7	.240***	-.025	.231***	.392***	.388***	.535***	-		
8	.373***	.268***	.417***	.672***	.665***	.494***	.601***	-	
9	.080	.284***	.132**	.384***	.397***	.544***	.515***	.550***	-

***=p<.001 **= p<.01 *= p<.05 †= p<.1
1=pitch harmonics, 2=formant transition section, 3=speech rate, 4=eloquence, 5=amiableness, 6=trust in a leader,   7=authenticity, 8=clarity, 9=absorption,

Table 3 shows the results of the linear regression analysis and Sobel test, which can prove the mediating effects of cognitive variables between pitch harmonics (independent) and trust in a leader (dependent). As **Table 3** suggests, pitch harmonics did not have a significant influence on trust in a leader in Step 1. In contrast, in Step 2-1, which inputs a mediating variable (eloquence), even pitch harmonics had a negative effect on trust in a leader, and eloquence had a significant positive effect on trust in a leader. Since pitch harmonics negatively influences trust in a leader, the Soble test value was negative, but eloquence positively affected trust in a leader. Similarly, amiableness, authenticity, and clarity had

mediating effects on trust in a leader. However, absorption did not have any significant mediating effect.

Table 3. The mediating effects of cognitive values between pitch harmonics and trust in a leader

		pitch harmonics	trust in leader	R^2 (Adj- R^2)	ΔR^2 (Adj- R^2)	F	Sobel test
step 1	β	.002		.000 (-.004)	-	.001	
	t	.035					
	sig	.972					
step 2-1 eloquence	β	-.181	.659	.401 (.395)	.401 (.399)	75.223***	-3.242***
	t	-3.368	12.266				
	sig	.001	.000				
step 2-2 amiableness	β	-.154	.605	.342 (.336)	.342 (.340)	58.348***	-2.661***
	t	-2.744	10.803				
	sig	.007	.000				
step 2-3 authenticity	β	-.134	.567	.303 (.297)	.303 (.303)	48.973***	-2.265**
	t	-2.332	9.897				
	sig	.021	.000				
step 2-4 clear	β	-.212	.573	.283 (.277)	.283 (.281)	44.381***	-3.258***
	t	-3.478	9.421				
	sig	.001	.000				
step 2-5 absorption	β	-.041	.547	.297 (.291)	.297 (.295)	47.592***	-0.732
	t	-.735	9.756				
	sig	.463	.000				
***=$p<.001$ **= $p<.01$ *= $p<.05$ †= $p<.1$							

Table 4 indicates the mediating effects of cognitive factors between formant transition section and trust in a leader. In these relationships, authenticity, clarity and absorption had significant mediating effects. However, eloquence and amiableness did not have mediating effects. Hence, we can find different effects of formant transition section from pitch harmonics.

Table 4. The mediating effects of cognitive values between formant transition section and trust in a leader

		formant transition section	trust in leader	R^2 (Adj- R^2)	ΔR^2 (Adj- R^2)	F	Sobel test
step 1	β	.020		000 (-.004)	-	.092	
	t	.304					
	sig	.762					
step 2-1 eloquence	β	-.060	.617	.374 (.368)	.374 (.372)	67.230***	-1.121
	t	-1.127	11.589				
	sig	.261	.000				
step 2-2 amiableness	β	-.078	.579	.325 (.319)	.325 (.323)	54.278***	-0.141
	t	-1.41 1	10.412				
	sig	.160	.000				
step 2-3 authenticity	β	.034	.536	.228 (.281)	.228 (.285)	45.412***	0.597†
	t	.597	9.523				
	sig	.551	.000				
step 2-4 clear	β	-.121	.527	.258 (.251)	.258 (.255)	39.096***	-1.980 *
	t	-2.02 7	8.836				
	sig	.044	.000				

step 2-5 absorption	β	-.146	.585	.315 (.309)	.315 (.313)	51.778***	-2.458**
	t	-2.537	10.17				
	sig	.012	.000				
***=p<.001 **= p<.01 *= p<.05 †= p<.1							

Table 5 shows the results of the linear regression analysis and Sobel test, which can prove mediating effects of cognitive variables between speech rate (independent) and trust in a leader (dependent). As Table 5 suggests, speech rate did not have a significant influence on trust in a leader in Step 1. In contrast, in Steps 2-1 to 2-5, which input mediating variables, all cognitive variables had positive effects on trust in a leader. Also, Soble test values supported their mediating effects between speech rate and on trust in a leader.

Table 5. The mediating effects of cognitive values between speech rate and trust in a leader

		Speech rate	trust in leader	R^2 (Adj- R^2)	ΔR^2 (Adj- R^2)	F	Sobel test
step 1	β	.006		.000 (-.004)	-	.008	
	t	.092					
	sig	.927					
step 2-1 eloquence	β	-.192	.666	.404 (.399)	.404 (.403)	76.312***	4.377***
	t	-3.564	12.354				
	sig	.000	.000				
step 2-2 amiableness	β	.-.169	.613	.346 (.340)	346 (.344)	59.444***	4.120***
	t	-3.004	10.903				
	sig	.003	.000				
step 2-3 authenticity	β	-.124	.564	.301 (.295)	.301 (.299)	48.443***	3.329***
	t	-2.164	9.842				
	sig	.032	.000				
step 2-4 clear	β	-.242	.595	.293 (.287)	.293 (291)	46.578***	5.366***
	t	-3.926	.9651				
	sig	.000	.000				
step 2-5 absorption	β	-.067	.552	.300 (.294)	.300 (.298)	48.206***	1.978*
	t	-1.186	9.818				
	sig	.237	.000				
***=p<.001 **= p<.01 *= p<.05 †= p<.1							

4. Conclusion

4.1 Results and Implications

This study demonstrated that a leader's voice characteristics have a relationship with their followers' psychology and cognition. The results of this study show that a leader's voice characteristics (pitch harmonics, formant transition section, and speech rate) do not directly increase trust in the leader. These features, however, affect the cognition of the leader's followers. Also, the mediating effects of cognitive variables (eloquence, amiableness, authenticity, clarity, and absorption) which increase trust in a leader are also important. The mediating effect of these cognitive variables was verified in the relationship of trust in leaders with all other voice characteristics, although the mediating effect of absorption between pitch harmonics and trust in leaders was not significant.

These results explain that the specific voice characteristics of a leader influence members' cognition through areas such as absorption and eloquence of speech. Thus, leaders need to

make an effort to utilise specific voice features. This is because voice characteristics can improve trust in a leader through the cognition of members. Leaders can gain trust in the content of their speeches through specific voice features, motivate more members, and improve their overall performance. Therefore leaders should strive to use an appropriate level of pitch harmonics, formant transition sections, and have the right speech rate when communicating with members.

In addition, the results of this research can be used to improve the performance of NLP. Few studies have been conducted on NLP with an agglutinative language. An inflectional language such as English and an agglutinative language differ in data analysis methods [40, 41]. This study was conducted on   Korean, an agglutinative language. As a result of the study, NLP parameters (pitch, velocity, formant) are presented in an agglutinative language system. NLP needs to focus on how humans recognize language [43, 44]. Understanding the parameters presented in this study will affect a NLP's response technology which can affect human cognition. Moreover, contextual factors need to be considered in NLP studies. When an NLP provides information to people, there may be more appropriate strategies for the situation, such as a clearer voice for giving information such as the news. Further, as AI develops, it will also require voices that can appeal or engage [45, 46]. Therefore, the input and output appropriate for each situation should be utilized in natural language processing. In addition, future research on other factors that may affect natural language (for example non-voice features such as pitch perturbation, energy contour, formant slope of f1 and f4, formant bandwidth change, average energy, etc.) will be necessary for the optimum use of NLP.

4.2 Limitations and Potential Areas of Future Study

This paper has some limitations, and we would also like to suggest several ideas for future studies. First, there are few samples in this paper, so   more samples of leaders and followers in a greater range of companies need to be collected and analysed . Then, greater individual (personality, gender and the like), corporate culture, and industrial differences will allow us to study the different influences of a leader's voice characteristics in greater depth. Second, there appear to be no relationships among several variables. Furthermore, we could not control all of the factors that clearly affected other variables, such as the contents of the speeches, or the pre-existing relationship between the followers and the leaders. Moreover, there exists other variables besides the five cognitive factors we looked at, and future research should be conducted to determine the relevance of these variables. Third, a person's voice characteristics can change. Therefore, it is also necessary to study how the perceptions of members changed after a leader changed one or more of their voice characteristics. Finally, we propose the need for research on the influence of voice features not only in regards to trust in a leader, but also other variables related to followers´ satisfaction, motivation, and performance. Finally, we need to continue to study what the voice of a better leader is. According to leadership contingency theory, there is not always one correct leadership trait or behaviour [47]. The leadership needed for each situation can be different. For example, charismatic leadership is needed in times of crisis, such as war or mergers and acquisitions; but self-leadership may be more appropriate to increase worker creativity. Hence, a leader needs to behave appropriately to each situation. Furthermore, according to leadership cognition theory, it is more important to consider how members recognize leaders' behaviours than what leaders do [6]. Individual members can recognize the actions of a leader differently. For instance, looking at a leader who comes to work early in the morning, one worker may recognize that the leader is sincere, but the other employee may perceive the leader to be strict. Hence, a leader needs to demonstrate behaviours or traits appropriate to each situation, and the members need to

acknowledge these behaviours and traits positively. Therefore, a good leader's voice should have characteristics appropriate to the situation. In a crisis situation, the leader will have to speak in a voice that can improve eloquence and absorption. In everyday situations, a voice that can increase amiableness and clarity will be more appropriate. Hence, future studies should try to find out what voice features are more suitable for various situations. Further, a discussion of good voices should be based on the cognition of members who listen to this voice. Therefore, as in this study, interdisciplinary studies such as sound engineering analyses and cognitive measurements of voice characteristics should continue.

References

[1] P. F. Drucker, Management Challenges for the 21st Century, NY, USA: HarperBusiness, 2001.

[2] G. A. Yukl, Leadership in organizations, NJ, USA: Pearson Education Inc, 2012.

[3] M. C. Rush and J. E. Russell, "Leader prototypes and prototype-contingent consensus in leader behavior descriptions," *Journal of Experimental Social Psychology*, vol. 24, no. 1, pp. 88-104, Jan. 1988. Article (CrossRef Link)

[4] R. R. Blake and J. S Mouton, The Managerial Grid, Houston, TX, USA: Gulf Publishing, 1994.

[5] S. A. Kirkpatick and E. A. Locke, "Leadership: do traits matter?," *The Executive*, vol. 5, no. 2, pp. 48-60, Feb. 1991. Article (CrossRefLink)

[6] J. R. Meindl, S. B. Ehrlich, and J. M. Dukerich, "The romance of leadership," *Administrative Science Quarterly,* vol. 30, no. 1, pp. 78-102, Mar. 1985. Aritcle (CrossRef Link)

[7] J. I. Lee, J. Y. Choi and H. K. Kang, "Analysis of Voice Quality Features and Their Contribution to Emotion Recognition," *Journal of Broadcast Engineering*, vol. 18, no. 5, pp. 771-774, Sep. 2013. Article (CrossRef Link)

[8] K. W. Choi and J. Y. Choi, "A study on emotion recognition and voice quality parameter systems," *Korean Institute of communications and Information Sciences*, pp. 825-828, Nov. 2008. Article (CrossRef Link)

[9] J. H. Seo, J. H. Sohn, and M. J. Bae, "A Study on the Voice Parameter by an Age Group," *Korean Institute of Communications and Information Sciences*, pp. 238-238, July 2004. Article (CrossRef Link)

[10] B. H. Yun, G. R. Baek, and M. J. Bae, "A Study on Building SVDB to Monitor a Soft Voice," *The Institute of Electronics and Information Engineers*, pp. 800-801, June 2011. Article (CrossRef Link)

[11] W. R. Jo and M. J. Bae, "On a Voice Color Change in the Fairy Tale System with Parent's Voice Color," *The Institute of Electronics and Information Engineers*, pp. 1163-1166, Nov. 1997. Article (CrossRef Link)

[12] E. H. Kim, K. H. Hyun, and Y. K. Kwak, "Noise Robust Emotion Recognition Feature: Frequency Range of Meaningful Signal," *Journal of the Korean Society for Precision Engineering*, vol. 23, no. 5, pp. 68-76, May 2006. Article (CrossRef Link)

[13] M. J. Bae and S. H. Lee, Digital Voice Analysis*,* Seoul, Korea: Dongyoung publish, 1998.

[14] H. W. Park, S. G. Bae, and M. J. Bae, "Analysis of Confidence and Control through Voice of Kim Jung-un," *International Information Institute,* vol. 19, no. 5, May 2016. Article (CrossRef Link)

[15] H. W. Park and S. W. Hahm, "A Study on Leaders' Voice and that Influences," *International Journal of Business Policy and Strategy Management*, vol. 3, no. 1, pp. 41-46, Dec. 2016. Article (CrossRef Link)

[16] K. T. Dirks and D. L. Ferrin, "Trust in leadership: meta-analytic findings and implications for research and practice," *Journal of Applied Psychology*, vol. 87, no. 4, pp. 611-628, Aug. 2002. Article (CrossRef Link)

[17] L. R. Rabiner and R. W. Schafer, "Introduction to Digital Speech Processing," *Foundations and Trends in Signal Processing*, pp. 1-194, 2007.

[18] L. R. Rabiner and R. W. Schafer, "Digital Speech Processing," *The Froehlich/Kent Encyclopedia of Telecommunications,* vol. 6, pp. 237-258, 2011.

[19] L. Rabiner and R. Schafer, Theory and Application of Digital Signal Processing, NY, USA: Pearson, 2010.

[20] L. Rabiner and R. Schafer, Digital Processing of Speech Signals, NY, USA: Pearson, 1978.

[21] W. C. Park, S. B. Lee, and S. H. Lee, Fundamentals of Sound Engineering, Korea: Chasong Press, 2009.

[22] W. Y. Yang, and Y. S. Jo, Digital Signal Processing and MATLAB, Seoul, Korea: Chungang University Press, 2001.

[23] Y. H. Song, J. H. Ahn, and M. J. Bae, "On the noise detection from correlation of near pitch waveforms," *GESTS Int'l Transactions on Computer Science and Engineering,* vol. 44, no. 1, pp. 45-54, 2008.

[24] H. W. Park, M. S. Kim, and M. J. Bae, " Improving Pitch Detection through Emphasized Harmonics in Time-Domain," in *Proc. of International Conferecne on Databae Theory and AplicationCommunications,* vol. 352, pp. 184-189, 2012. Article (CrossRef Link)

[25] W. R. Jo, S. Y. Choi, and M. J. Bae, "A Study on the Pitch Search Time Reduction of G.723.1 Vocoder by Improved Hybrid Domain Cross-correlation," *The Transactions of The Korean Instititute of Electrical Engineers,* vol. 59, no. 12, pp. 2324-2328, Dec. 2010. Article (CrossRef Link)

[26] S. M. Lee, "Paralinguistic Communication Analysis of Popular AM Radio Talk-show Personality," *Speech Research*, vol. 1, pp. 157-158, 1999.

[27] C. Morton, Change Your Voice, Change Your Life, NY, USA: Macmillan Publishing Co., 1984.

[28] M. Morishima, "Information sharing and firm performance in Japan," *Industrial Relations: A Journal of Economy and Society*, vol. 30, no. 1, pp. 37-61, Jan. 1991. Article (CrossRef Link)

[29] R. L. Street Jr., R. M. Brady, and R. Lee, "Evaluative responses to communicators: The effects of speech rate, sex, and interaction context," *Western Journal of Speech Communication*, vol. 48, no. 1, pp. 14-27, June 2009. Article (CrossRef Link)

[30] C. J. Jeong and M. J. Bae, "Robust Pitch Analysis and Detection in Noise Environment," *Korean Institute of Communications and Information Sciences*, pp. 1330-1333, July 2003. Article (CrossRef Link)

[31] B. Lewis, The Technique of Television Announcing, NY, USA: Hastings House, 1966.

[32] A. Mehrabian and M. Williams, "Nonverbal concomitants of perceived and intended persuasiveness," *Journal of Personality and Social Psychology*, vol. 13, no. 1, pp. 37-58, 1969. Article (CrossRef Link)

[33] D. M. Rousseau, S. B. Sitkin, R. S. Burt, and C. Camerer, "Not so different after all: A cross-discipline view of trust," *Academy of Management Review*, vol. 23, no. 3, pp. 393-404, July 1998. Article (CrossRef Link)

[34] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An integrative model of organizational trust," *Academy of Management Review*, vol. 20, no. 3, pp. 709-734, July 1995. Article (CrossRef Link)

[35] J. B. Rotter, "A new scale for the measurement of interpersonal trust," *Journal of Personality*, vol. 35, no. 4, pp. 651-665, Dec. 1967. Article (CrossRef Link)

[36] P. M. Podsakoff, S. B. MacKenzie, R. H. Moorman, and R. Fetter, "Transformational leader behaviors and their effects on followers' trust in leader, satisfaction, and organizational citizenship behaviors," *The Leadership Quarterly*, vol. 1, no. 2, pp. 107-142, 1990. Article (CrossRef Link)

[37] T. Nasukawa and J. H. Yi, "Sentiment analysis: Capturing favorability using natural language processing," in *Proc. of the 2nd international conference on Knowledge capture,* pp. 70-77, Oct. 2003. Article (CrossRef Link)

[38] E. Loper and S. Bird, "NLTK: the natural language toolkit," in *Proc. of the ACL 2004 on Interactive poster and demonstration sessions,* May 2004. Article (CrossRef Link)

[39] E. Cambria and B. White, "Jumping NLP curves: A review of natural language processing research," *IEEE Computational Intelligence Magazine,* vol. 9, no. 2, pp. 48-57, Apr. 2014. Article (CrossRef Link)

[40] M. W. Han, S. C. Park, H. B. Lee, J. H. Yeon, and S.G. Lee, "Natural language processing on Korean language : A survey," in *Proc. of The Korean Institute of Information Scientists and Engineers,* pp. 681-683, June 2015. Article (CrossRef Link)

[41] H. J. Lee and J. W. Kim, "A Study on the Natural Language Processing(NLP) Technical and Standardization Trend," in *Proc. of Symposium of the Korean Institute of communications and Information Sciences*, pp. 876-877, June 2017. [Article (CrossRef Link)](#)

[42] M. Kienast and W. F. Sendlmeier, "Acoustical analysis of spectral and temporal changes in emotional speech," *ISCA Tutorial and Research Workshop (ITRW) on Speech and Emotion,* Sep. 2000. [Article (CrossRef Link)](#)

[43] W. E. Zhang, Q. Z. Sheng, A. Alhazmi, and C. Li, "Adversarial attacks on deep-learning models in natural language processing: A survey,*" ACM Transactions on Intelligent Systems and Technology,* vol. 11, no. 3, pp. 1-41, Apr. 2020. [Article (CrossRef Link)](#)

[44] J. Eisenstein, Introduction to natural language processing, MA, USA: The MIT Press, 2019.

[45] C. Edwards, A. Edwards, B. Stoll, X. Lin, and N. Massey, "Evaluations of an artificial intelligence instructor's voice: Social Identity Theory in human-robot interactions," *Computers in Human Behavior*, vol. 90, pp.357-362, Jan. 2019. [Article (CrossRef Link)](#)

[46] C. Sun, Z. J. Shi, X. Liu, A. Ghose, X. Li, and F. Xiong, "The Effect of Voice AI on Consumer Purchase and Search Behavior," *NYU Stern School of Business*, pp. 1-43, Oct, 2019. [Article (CrossRef Link)](#)

[47] P. G. Northouse, Leadership: Theory and practice, Newbury Park, CA, USA: SAGE Publications, 2018.

Sang Woo Hahm is currently a professor of Business Administration at Semyung University, Korea. He received his PhD, MS and BS in School of Business Administration from Soonsil University. His current research interests include e-business, management, The fourth industrial revolution, organizational behavior, leadership, interdisciplinary research and the like.

Hyung Woo Park received a Ph.D., an M.S., and a B.S. in Electrical Engineering from Soongsil University. He is a Staff Researcher at the Quantum Dot Display Commerialization Team at Samsung Display, Korea. His current research interest includes Sound signal processing, Big data analysis, Voice analysis, Noise reduction system, Wave field synthesis, Advanced acoustic radiation system, Railway noise, and Internet of Things.