The organizational voice: The importance of voice pitch and speech rate in organizational crisis communication

Aurélie De Waele and An-Sofie Claeys
KU Leuven
Verolien Cauberghe
Ghent University

Author note:
Aurélie De Waele is a PhD student at the Institute for Media Studies at KU Leuven (Belgium). An-Sofie Claeys is assistant professor in Corporate Communication at the Institute for Media Studies at KU Leuven (Belgium). Verolien Cauberghe is associate professor in Communication Management at Ghent University (Belgium).

Acknowledgements:
This research was supported in part by a grant from the Agency for Innovation by Science and Technology in Flanders (IWT).

Contact:
Aurélie De Waele (aurelie.dewaele@soc.kuleuven.be) or An-Sofie Claeys (ansofie.claeys@soc.kuleuven.be), KU Leuven, Institute for Media Studies, Parkstraat 45, 3000 Leuven, Belgium.
Verolien Cauberghe (verolien.cauberghe@ugent.be), Ghent University, Department of Communication Sciences, Korte Meer 7-9-11, 9000 Gent, Belgium.
Abstract

Prior research has mainly focused on verbal aspects of crisis communication (i.e., crisis response strategies) (e.g., Coombs, 2007). However, crisis communication is frequently spread through audiovisual messages, which also contain nonverbal cues (Coombs & Holladay, 2009). Such vocal (e.g., voice pitch) and visual (e.g., facial expression) cues affect listeners’ perceptions of a speaker (e.g., Tigue et al., 2012). Yet only a few studies have investigated visual cues in crisis communication (e.g., Arpan, 2002; Gorn, Jiang, & Venkataramani, 2008), while research on vocal cues is almost nonexistent (Claeys & Cauberghe, 2014). Therefore, this study examines the impact of two vocal cues, namely pitch and speech rate, in combination with the crisis communication strategy on post-crisis reputation, by means of a 2 (voice pitch: low - high) × 2 (speech rate: slow - fast) × 2 (crisis response strategy: deny - rebuild) between-subjects experimental design, involving a fictitious preventable crisis. Results show that voice pitch and speech rate affect post-crisis reputation, but only in case of a matched (i.e., rebuild) crisis response strategy. Moreover, the interaction between verbal and vocal cues is mediated by attractiveness of the voice and by empathy towards the spokesperson.
Public relations research is a rapidly expanding field. One of the most researched themes in public relations is crisis communication (Avery, Lariscy, Kim, & Hocke, 2010). Crisis communication research and theory building have mainly focused on verbal aspects of crisis communication, and more specifically on the effects of crisis response strategies (e.g., apologizing, denial) (e.g., Coombs, 2007) on post-crisis reputation. However, corporate communication becomes increasingly audiovisual, mainly because of the frequent use of social media, which includes video messages (e.g., YouTube). Those audiovisual messages convey not only verbal, but nonverbal cues as well (Coombs & Holladay, 2009). Research in aligned fields, such as political communication and marketing communication, shows that such nonverbal cues can influence an audience’s perception of a communicator (e.g., Tigue, Borak, O’Connor, Schandl, & Feinberg, 2012; Chattopadhyay, Dahl, Ritchie, & Shahin, 2003). Moreover, several studies have indicated that those nonverbal cues can interact with the content of the message (e.g., Chattopadhyay et al., 2003). Despite the potential importance of nonverbal communication for organizational spokespersons, however, there has been limited research on the effects of nonverbal cues in crisis communication.

Some studies have examined the effects of visual cues in crisis communication (e.g., Claeys & Cauberghe, 2014a; Hong & Len-Ríos, 2015). However, the research on vocal cues in crisis communication is limited to the effect of only one vocal cue (voice pitch) (Claeys & Cauberghe, 2014a). Moreover, we only know how voice pitch affects perceptions of a spokesperson (Claeys & Cauberghe, 2014a) but not how this spills over on perceptions of the organization itself. Therefore, this study explores how two vocal cues, namely pitch and speech rate, affect the evaluation of an organization in crisis and how. Moreover, because vocal and verbal cues interact in aligned fields, such as marketing communication (e.g., Chattopadhyay et al., 2003), this study examines the interplay between verbal (i.e., crisis response strategy) and vocal cues (i.e., voice pitch and speech rate) in times of organizational crisis. The findings of this study will expand the field of crisis communication research by exploring the impact of vocal cues on post-crisis reputation and their interaction with the more researched area of crisis response strategies. On the one hand, this is relevant for communication research, because the findings will show if the impact of vocal cues should be taken into account in researching crisis communication outcome. On the other hand, this is relevant for crisis communication practice, because the findings will indicate how practitioners should take into account the speech characteristics of a spokesperson in times of crisis.
Verbal aspects of crisis communication

Crisis communication research has mainly focused on the content of crisis communication. Particularly, most studies have examined the impact of crisis response strategies, such as apologizing or denying responsibility, on organizational reputation (Avery et al., 2010). The most important theory regarding crisis response strategies is Situational Crisis Communication Theory (SCCT) (Coombs, 2007). SCCT explores how the characteristics of the crisis situation (e.g., crisis responsibility, crisis severity) affect the impact of crisis response strategies, and it provides guidelines about which crisis response strategy is best suited under which conditions. According to SCCT, the reputation threat of a crisis is mainly determined by crisis responsibility attributions by the public (Coombs & Holladay, 2002). Therefore, SCCT distinguishes three crisis type clusters, based on attributions of responsibility, namely victim, accidental and preventable crises, increasing in the amount of crisis responsibility attributions and reputational damage. According to SCCT, an organization in crisis should accept as much responsibility for the events as the public attributes (Coombs & Holladay, 2002). Therefore, there are also three clusters of crisis response strategies, differing in the extent to which the organization takes responsibility for the crisis. Deny strategies disprove the organization’s responsibility for the event. Diminish strategies accept involvement in the crisis, but try to minimize responsibility for it. Rebuild strategies accept responsibility for the crisis and attend to victim concerns (Coombs, 2006). Consequently, SCCT matches deny strategies to victim crises, diminish strategies to accidental crises, and rebuild strategies to preventable crises (Coombs, 2007). Some experimental studies confirm the assumption that a crisis response that is matched to the amount of responsibility attributions leads to less reputation damage (e.g., Huang, 2006). In particular, several studies have indicated that organizations confronted with a preventable crisis should take their responsibility for the events by means of an appropriate rebuild strategy (Coombs, 1995; 2007; Coombs & Holladay, 1996; 2002). This leads to the following hypothesis:

H1: A rebuild crisis response strategy leads to a better post-crisis reputation in the case of a preventable crisis than a deny crisis response strategy.

In recent years, studies on the content of crisis communication are being complemented with insights on other elements of the crisis message, such as timing and framing (Arpan & Roskos-Ewoldsen, 2005; Claeys & Cauberghe, 2012; Claeys, Cauberghe, & Leysen, 2013; van der Meer & Verhoeven, 2014).

However, people frequently receive news about organizations via audiovisual media, such as television and video messages on the Internet. Also, organizations increasingly use social
media for crisis communication (Schultz, Utz, & Göritz, 2011; Utz, Schultz, & Glocka, 2013), which often includes video messages as well (e.g., YouTube). Crisis communication thus becomes increasingly audiovisual. Such audiovisual crisis communication not only conveys the content of a message, but also visual and vocal cues (i.e., nonverbal communication) (Coombs & Holladay, 2009). Several authors have recognized the importance of nonverbal communication in times of crisis, for example in order to express emotions (van der Meer & Verhoeven, 2014). Yet, nonverbal communication is an aspect that remains relatively unexplored in crisis communication research.

Nonverbal communication

Nonverbal communication has been studied extensively in the context of impression formation (e.g., Tigue et al., 2012), because people’s opinion about another person is not only based on what a person says (verbal cues), but also on visual and vocal cues (Sporer & Schwandt, 2006). Some insights of research on visual cues in impression formation, have been applied to crisis communication research, by investigating the impact of visual cues on the evaluation of the crisis and the organization in crisis. The effects of a spokesperson’s ethnicity (Arpan, 2002), race (Hong & Len-Riós, 2015), the degree to which the spokesperson has a babyface (Gorn, Jiang, & Johar, 2008), and the use of powerful gestures (Claeys & Cauberghe, 2014a) on the evaluation of the organization in crisis have been investigated.

Besides research on visual nonverbal cues, there is an extensive amount of research on vocal cues in impression formation as well. According to social psychological and interpersonal communication research, vocal cues have a significant influence on the perception of a person (e.g., Imhof, 2010). People link vocal cues to certain personality traits, regardless of whether the speaker effectively possesses these characteristics (i.e., voice stereotypy) (Aronovitch, 1976). Two of the most important vocal cues in the context of impression formation are voice pitch and speech rate (e.g., Apple, Streeter, & Krauss, 1979).

First, pitch is typically linked to perceptions of dominance. Several studies that have manipulated pitch experimentally, show that a lower pitched voice leads to the perception of more physical and social dominance, attractiveness and leadership of the speaker (e.g., Puts, Gaulin, & Verdolini, 2006; Tigue et al., 2012), while a higher pitched voice scores low on dominance ratings and comes across as nervous and emotionally instable (Apple et al., 1979; Imhof, 2010). Second, voice pitch has also been connected to perceptions of credibility. A voice with a lower pitch is perceived as more honest, trustworthy and credible (Apple et al., 1979; Tigue et al., 2012), while a higher pitch is linked to deception (Apple et al., 1979; DePaulo et al., 2003; Sporer & Schwandt, 2006). Third, voice pitch is associated with
perceptions of competence. A lower pitched voice leads to a better perception of a speaker in terms of intelligence, expertise, and competence (Tigue et al., 2012). A fourth aspect that is linked to vocal pitch, is perceived empathy of the speaker. A study investigating vocal cues in doctor-patient communication, showed that a lower pitch is perceived as more caring and sympathetic (McHenry, Parker, Baile, & Lenzi, 2012).

The findings of these studies concerning pitch in impression formation, have been applied to several domains, such as political communication and advertising research. These studies show that pitch is not only linked to personality judgments, but also impacts the attitudes and behavior of the listener towards the speaker. For example, two studies investigated the effect of a politician’s voice pitch on listeners’ voting behavior, by manipulating the politician’s pitch experimentally. In both studies, participants chose to vote for the politician with the lower pitch significantly more often (Klofstad, Anderson, & Peters, 2012; Tigue et al., 2012). A different study, in the domain of advertising research, showed that radio advertisements with a spokesperson with a low pitch led to better attitudes towards the spokesperson and a better advertisement effectiveness than a spokesperson with a high pitch (Martín-Santana, Muela-Molina, Reinares-Lara, & Rodríguez-Guerra, 2015).

Besides vocal pitch, there is another vocal cue that has an important influence on the perception of the speaker, namely speech rate (e.g., Ray, 1986). Several studies on personality perception have linked a fast speech rate with the perception of a competent speaker (e.g., Brown, Strong, & Rencher, 1974; Ray, 1986; Street & Brady, 1982). Also, a fast speech rate can function as a credibility cue, which enhances persuasiveness (Miller, Maruyama, Beaber, & Valone, 1976). This has led researchers to investigate the effects of speech rate in sales and advertising contexts. A study investigating the relationship between vocal cues and sales performance in salespeople, showed that a fast speech rate is related to an increase in sales performance (Peterson, Cannito, & Brown, 1995). In advertising research, several studies have shown that a fast speech rate is more persuasive and leads to more positive consumers’ attitudes (e.g., Skinner, Robinson, Robinson, Sterling, & Goodman, 1999).

However, the impact of vocal cues, such as pitch and speech rate, on listeners’ perceptions and attitudes towards the speaker, is not only based on their link with certain personality traits. Listeners’ judgments of a speaker are also based on the attractiveness of the speaker’s voice. Literature indicates a “vocal attractiveness stereotype” (Zuckerman & Driver, 1989), which means that vocal attractiveness elicits a halo effect (Surawski & Ossoff, 2006). Consequently, attractive voices lead to more positive personality judgments (Surawski & Ossoff, 2006; Tigue et al., 2012; Zuckerman & Driver, 1989). Thus, the effects of vocal cues on listeners’
perceptions and attitudes towards the speaker often occur as a function of vocal attractiveness (Zuckerman & Driver, 1989).

Finally, besides their impact on impression formation, vocal cues also play a role in conveying emotions (e.g., Banse & Scherer, 1996). On the one hand, they play a role in expressing emotions. Several studies have listed which vocal cues are associated with which emotions, and pitch and speech rate seem to be two important cues for expressing emotions, such as anger, happiness, sadness or fear (e.g., Murray & Arnott, 1993). On the other hand, vocal cues also play a role in the perception of emotions. Pitch and speech rate are important cues that are used by a listener to deduce emotions from a speaker’s voice (e.g., Banse & Scherer, 1996; Johnstone & Scherer, 2000). The findings that the voice both expresses emotions and is used by listeners to infer a speaker’s emotions, has led researchers to explore whether the expression of emotions in the voice has an effect on the emotional state of the listener (e.g., Hawk, Fischer, Van Kleef, & Gerben, 2012). Research indicates that emotions in the voice indeed affect the emotional state of the listener, in that sense that listeners have an emotional reaction towards the observed emotions in the speaker’s voice (Hawk et al., 2012). Thus, vocal cues, such as pitch and speech rate, also seem to play a role in affecting the emotions of the listener towards the speaker.

**Vocal cues in crisis communication**

Although the effects of pitch and speech rate (e.g., competence, credibility) that have been found in aligned fields, such as political communication and marketing communication (e.g., Tigue et al., 2012), are also relevant in crisis communication, research on the influence of vocal cues in crisis communication is very limited. One study has investigated the effect of one vocal cue on the perception of an organizational spokesperson in times of crisis, by manipulating the voice pitch of the spokesperson (Claeys & Cauberghe, 2014a). The results showed that a spokesperson who speaks with a low pitch during a crisis is perceived as more competent than a spokesperson with a high pitch. Moreover, the results of the experiment illustrated that the effect of voice pitch on competence was mediated by perceptions of powerfulness. However, the impact of vocal cues on the perception of the organization (i.e., post-crisis reputation) has not been investigated before.

In addition, another important vocal cue, namely speech rate, has not yet been examined in times of crisis. It is not clear if the effect of pitch and speech rate in crisis communication would be the sum of the separate effects of those cues, or if an interaction between those cues would occur. Some studies have suggested that a fast speech rate leads people to pay more attention to other vocal cues, such as pitch (e.g., Chattopadhyay et al., 2003). Other studies
have indicated that there is a natural covariation between pitch and speech rate (e.g., Black, 1961), which could make speech with levels of pitch and speech rate that do not naturally covary, sound odd. It is thus not clear if and how pitch and speech rate affect crisis communication outcome. This leads to the following research question:

**RQ1: How do the vocal cues pitch and speech rate of an organizational spokesperson affect post-crisis reputation?**

It is not only important to investigate the impact of the vocal cues pitch and speech rate in crisis communication, but also to investigate how this possible impact could be explained. The study of Claeys & Cauberghe (2014a) has shown that voice pitch affects the perception of a spokesperson’s competence through perceptions of powerfulness. However, research in aligned fields (e.g., political communication) has shown that the impact of vocal cues could also be explained through vocal attractiveness (e.g., Tigue et al., 2012). Voices that are rated as more attractive, lead to more positive judgments of a spokesperson (e.g., Surawski & Ossoff, 2006; Tigue et al., 2012; Zuckerman & Driver, 1989). This leads to the question if the possible effects of vocal cues on post-crisis reputation in crisis communication, could be mediated by the attractiveness of the organizational spokesperson’s voice.

**RQ2: Is the impact of vocal cues on post-crisis reputation mediated by the attractiveness of the spokesperson’s voice?**

Furthermore, literature on crisis communication shows that emotions have an impact on crisis communication outcome (Claeys & Cauberghe, 2014b; van der Meer & Verhoeven, 2014). Those emotions are not only communicated by the content of the crisis response, but particularly by a spokesperson’s voice (van der Meer & Verhoeven, 2014). Research from social psychology has shown that vocal cues, such as pitch and speech rate, indeed express emotions (e.g., Murray & Arnott, 1993). Those emotions in the voice can affect the emotional state of the listener by evoking an emotional reaction towards the speaker’s observed emotions (Hawk et al., 2012). Because affective empathy towards a person arises from such an emotional reaction towards the observation of another person’s emotions (Lawrence, Shaw, Baker, Baron-Cohen, & David, 2004), it is possible that vocal cues can thus elicit empathy towards the speaker. Therefore, the question rises whether the impact of pitch and speech rate in times of crisis can be explained by empathy towards the spokesperson. This leads to the following research question:

**RQ3: Is the impact of vocal cues on post-crisis reputation mediated by empathy?**
The interplay between verbal and vocal cues

In communication, verbal and nonverbal cues are mostly combined. Therefore, several authors have suggested that verbal and nonverbal behavior should not be studied separately, but needs an integrated approach (Jones & LeBaron, 2002). Consequently, the interaction between verbal and nonverbal cues is an extensively studied topic in several areas of communication research, and specifically in persuasive communication research (e.g., Gelinas-Chebat & Chebat, 1992).

Such studies on the interaction between vocal cues and the content of a message in persuasive communication have shown different results. One view on the interaction between nonverbal (e.g., vocal) cues and verbal cues suggests that communication is more persuasive if the verbal and nonverbal cues of a message are congruent (Johnson, McCarty, & Allen, 1976). However, other studies indicate that the interaction between vocal cues and content depends on the involvement of the receiver of the message. They suggest that in the case of a highly involved receiver, the receiver is persuaded by the content of the message, such as the quality of the arguments. In the case of a low level of involvement, the receiver relies on peripheral cues, (e.g., vocal cues), to form their opinion (e.g. Gelinas-Chebat & Chebat, 1992; Reinhard & Sporer, 2010). However, other studies show that nonverbal cues themselves can impact to which extent persuasion occurs through the content of the message or through other nonverbal cues. For example, in advertising research, a fast speech rate undermines the extent to which people process the ad and therefore decreases the effects of the content of the message, while increasing the effects of peripheral cues, such as pitch (Chattopadhyay et al., 2003; Megehee, Dobie, & Grant, 2003; Moore, Hausknecht, & Thamodaran, 1986). It is thus clear that there is an interaction between vocal cues and the content of the message, but the nature of this interaction is indistinct.

In crisis communication, the interaction between vocal cues and the content of the crisis response has not been investigated before. Nevertheless, spokespersons send out both verbal (e.g., crisis response strategy) and vocal cues (e.g., pitch and speech rate) together when communicating in television or radio interviews or video messages. This leads to the question whether the effects of a spokesperson’s vocal cues in times of crisis are influenced by the used crisis response strategy (i.e., a matched (rebuild) strategy or a mismatched (deny) strategy):

*RQ4: What is the interplay between an organizational spokesperson's verbal and vocal cues on post-crisis reputation in times of crisis?*
Method

Design and Stimuli

The hypothesis and research questions were examined by means of a 2 (voice pitch: low vs. high) × 2 (speech rate: slow vs. fast) × 2 (crisis response strategy: match vs. mismatch) between-subjects experimental design. Participants read a text, describing a fictitious organizational crisis and then listened to an audio recording with a fictitious crisis response that manipulated each of the three independent variables. As such, each participant randomly received one of eight conditions.

The crisis situation that was described in the text, occurred at a fictitious company, in order to prevent any confounding effects of an existing pre-crisis reputation (Laufer & Jung, 2010). The text described a preventable crisis, which is a crisis type with high responsibility attributions that leads to the highest level of reputation threat (Coombs, 2007). Moreover, the crisis scenario was chosen to elicit moderate levels of crisis involvement, which is important to investigate the effects of both central (i.e., crisis response strategy) and peripheral cues (i.e., vocal cues) of a message (Petty & Cacioppo, 1984). Because personal relevance of the crisis increases involvement (Petty, Cacioppo, & Schumann, 1983; Smith & Shaffer, 1995), the scenario described a crisis concerning everyday products (which were highly relevant to the participants), but was set in an area of the country that was distant and culturally different from where the participants lived, which lowered personal relevance.

In addition, to increase the external validity of this study and as such the generalizability of the findings, two different crisis scenarios were used. Each scenario was, however, similar with regard to perceptions of organizational responsibility, crisis severity and crisis involvement. In the first scenario, people got severely ill after eating yoghurt that was contaminated with listeria. The cause of the contamination was poor hygiene at a dairy company. In the second scenario, people got severely ill after using toothpaste that contained a toxic chemical. The chemical ended up in the toothpaste as a result of deficient control mechanisms at a company producing care products.

After reading the text, participants listened to an audio recording from a fictitious radio interview given by the CEO of the organization in crisis. We recorded the voice of a male person with a normal pitch (106 Hz) (Skuk & Schweinberger, 2014) and speech rate (4.32 syllables/second) (e.g., Chattopadhyay et al., 2003). The person read an adapted crisis response for each of both crisis scenarios. Crisis response strategy was manipulated by recording two versions of the crisis response for each scenario. One version contained a rebuild crisis response strategy (i.e., the CEO stated that his company took full responsibility for the events), which
was matched to the crisis type (preventable crisis), according to SCCT (Coombs, 2007). The other version contained a deny crisis response strategy (i.e., the CEO stated that his company denied all responsibility for the events), which is considered a mismatched strategy in the case of a preventable crisis. All original recordings were similar in length, content, average pitch and speech rate. The only difference was the crisis response strategy used and the crisis situation that was addressed, resulting in two recordings for each scenario.

In order to manipulate voice pitch and speech rate, the audio fragments were then edited by means of the computer software Praat (cf. Boersma, 2001). This software allowed us to manipulate pitch and speech rate separately, while keeping all other factors constant (e.g., intensity, intonation) (Imhof, 2010). Voice pitch was manipulated by raising and lowering the baseline level with 20% (e.g., Apple et al., 1979). Speech rate was manipulated by increasing and decreasing the baseline level with 15%, which was in this case the boundary value for still sounding natural. This resulted in eight different audio conditions for each crisis scenario.

Participants and Procedure

A sample of 241 respondents that was representative of the Dutch-speaking Belgian population in terms of age, gender and sociodemographic variables, participated in the study. Participants were recruited via a market research bureau, which invited them to participate in an online survey. 14 respondents who indicated having hearing problems were omitted from the dataset. Hence, 227 participants were included in further analysis. Participants were divided randomly across the crisis scenarios and experimental conditions by the survey website. They first read the text describing the crisis scenario and then they listened to an audio fragment with the crisis response of the CEO. Finally, they filled out the online questionnaire containing manipulation checks, measures of the dependent variables and measures of sociodemographic variables. The average age of the participants was 51 years ($SD = 14.28$; range = 20 - 79). Approximately 46% of them were male and 54% were female.

Measures

The questionnaire first contained manipulation checks. Then, participants filled in questions regarding the dependent and mediating variables.

Crisis response strategy. In order to check for the manipulation of crisis response strategy, participants rated the reaction of the CEO on a one-item seven-point semantic differential scale ranging from “denied all crisis responsibility” to “took full responsibility for the crisis”. This scale is related to the definition of crisis response strategies from SCCT (Coombs, 2007), which classifies crisis response strategies according to the amount of responsibility they take (Claeys & Cauberghe, 2014b).
**Pitch.** The manipulation of pitch was checked by measuring the average pitch of the audio fragments. Average pitch was measured with the software Praat (Analyse periodicity: to pitch - Query: Get mean) (cf. Boersma, 2001).

**Speech rate.** The manipulation of speech rate was checked by measuring the average speech rate of the audio fragments. Speech rate was measured by counting the syllables in the audio fragment and dividing that number by the number of seconds of speech (e.g., Chattopadhyay et al., 2003).

**Organizational responsibility.** In order to check if the preventable crisis scenario, which should inflict high attributions of responsibility (Coombs & Holladay, 2002), was effectively perceived as a preventable crisis, perceptions of organizational responsibility were measured. A two-item seven-point Likert scale was used from Griffin, Babin, & Darden (1992), measuring the organizational responsibility for a crisis (e.g., “How responsible was the organization itself with respect to the crisis?”) ($\alpha = .90$).

**Involvement and severity.** Crisis involvement and severity were measured to check for differences between the two crisis scenarios. Involvement was measured by a three-item seven-point Likert scale ranging from 1 (completely disagree) to 7 (completely agree) (Kopalle & Lehmann, 2001) ($\alpha = .91$) (e.g. “These events are very important to me”). In order to measure perceived crisis severity, participants rated the crisis events on a three-item 7-point semantic differential scale from Arpan & Pompper (2003) (e.g., “not at all severe - extremely severe”) ($\alpha = .95$).

**Organizational reputation.** Organizational reputation was measured as a dependent variable using items from the reputation quotient of Fombrun, Gardberg and Sever (2000). The reputation quotient measures organizational reputation on an emotional and a rational dimension. Because the experiment involved a fictitious crisis situation, some items from the rational dimension were difficult to rate because of a lack of background information about the organization (e.g., “This organization develops innovative products and services”). Therefore, a short version of the reputation quotient was used, consisting of nine 7-point disagree-agree Likert items ($\alpha = .95$) (e.g., “I have a good feeling about this company”).

**Vocal attractiveness.** Attractiveness of the CEO’s voice was measured using a two-item 7-point Likert scale from Surawski & Ossoff (2006), on which participants indicated to which extent they found the voice “attractive” and “likeable” ($\alpha = .88$).

**Empathy.** Empathy towards the CEO was measured using an eight-item 7-point Likert scale, which assessed participants’ feelings towards the CEO (e.g., sympathy, concern, compassion) ($\alpha = .94$) (McCullough, Fincham, & Tsang, 2003).
Results

Pretest

A pretest was conducted, in order to check for any confounding effects of unintended differences between the crisis scenarios, and in order to measure if the crisis scenarios were indeed perceived as preventable crises with moderate levels of involvement. An independent t-test showed that there were no significant differences between the two scenarios in terms of crisis responsibility \((t(76) = -1.41, p = .16)\) and that both crises elicited high levels of perceived crisis responsibility (i.e. preventable crisis) \((M_{\text{scenario1}} = 5.78, SD = 1.17; M_{\text{scenario2}} = 6.13, SD = 1.02)\). In addition, the scenarios did not differ in terms of the involvement each crisis elicited \((t(76) = -.48, p = .63)\) and both events resulted in moderate levels of involvement \((M_{\text{scenario1}} = 4.00, SD = 1.47; M_{\text{scenario2}} = 4.15, SD = 1.37)\). Finally, the scenarios did not differ in terms of perceived crisis severity \((t(76) = -.49, p = .62; M_{\text{scenario1}} = 5.25, SD = 1.45 \text{ vs. } M_{\text{scenario2}} = 5.41, SD = 1.32)\).

Manipulation checks

The manipulation of each independent variable was examined in the experimental study. An independent samples t-test tested the manipulation of the first independent variable, namely crisis response strategy. The test showed that in the case of a rebuild crisis response strategy \((M = 6.13, SD = 1.32)\), the organization was perceived as taking significantly more responsibility for the crisis \((t(225) = -19.18, p < .001)\) than in the case of a deny crisis response strategy \((M = 2.12, SD = 1.75)\). Thus, the manipulation of the verbal crisis communication was successful.

The manipulation of the two vocal cues was checked by measuring the average pitch and speech rate of the audio fragments. The average pitch of the fragments with increased pitch was 125 Hz, which is approximately 20 % higher than the average baseline level of 106 Hz. The average pitch of the fragments with decreased pitch was 84 Hz, which is approximately 20 % lower than the average baseline level of 106 Hz. The average speech rate of the fragments with fastened speech rate was 5.09 syllables per second, which is approximately 15 % higher than the average baseline level of 4.32 syllables per second. The average speech rate of the fragments with slowed speech rate was 3.76 syllables per second, which is approximately 15 % lower than the average baseline level of 4.32 syllables per second. The manipulations of pitch and speech rate were thus as intended.

Finally, an independent t-test confirmed that there were no significant differences between both crisis scenarios in terms of crisis responsibility \((t(225) = -.55, p = .59)\) and that both crises elicited high levels of perceived crisis responsibility (i.e., preventable crisis) \((M_{\text{scenario1}} = 5.52,\)
$SD = 1.50$; $M_{\text{scenario2}} = 5.63$, $SD = 1.37$). In addition, the scenarios did not differ in terms of the involvement each crisis elicited ($t(225) = -0.04$, $p = .97$) and both events resulted in moderate levels of involvement ($M_{\text{scenario1}} = 4.11$, $SD = 1.63$; $M_{\text{scenario2}} = 4.12$, $SD = 1.58$). Finally, the scenarios did not differ in terms of perceived crisis severity ($t(225) = -0.35$, $p = .73$; $M_{\text{scenario1}} = 5.56$, $SD = 1.35$ vs. $M_{\text{scenario2}} = 5.62$, $SD = 1.29$).

**Hypotheses and research questions**

To test the hypothesis and research question 1 and 4, a three-way ANOVA was conducted. First, there was no main effect of crisis response strategy on post-crisis reputation ($F(1, 219) = .00$, $p = .99$). The use of a matched crisis response strategy (i.e., rebuild) did not result in a significantly better post-crisis reputation ($M_{\text{match}} = 3.11$, $SD = .12$) than a mismatched crisis response strategy (i.e., deny) ($M_{\text{mismatch}} = 3.11$, $SD = .13$). As such, hypothesis 1 was not confirmed.

The three-way ANOVA also addressed research questions 1 and 4. Regarding the first research question (How do the vocal cues pitch and speech rate of an organizational spokesperson affect post-crisis reputation?), there was no main effect of pitch ($F(1, 219) = .06$, $p = .80$), nor of speech rate ($F(1, 219) = .23$, $p = .64$).

The analysis did, however, show a significant three-way interaction between crisis response strategy, pitch and speech rate on post-crisis reputation ($F(1, 219) = 10.27$, $p < .01$), which offers insight into the fourth research question (What is the interplay between an organizational spokesperson’s verbal and vocal cues on post-crisis reputation in times of crisis?). Two follow-up two-way ANOVAs were conducted to look at this interaction in detail and examine how the crisis response strategy affects the impact of each vocal cue. A first analysis showed that in the case of a mismatched crisis response strategy (i.e., deny), there were no effects of the vocal cues pitch ($F(1, 99) = .18$, $p = .68$) and speech rate ($F(1, 99) = .05$, $p = .83$) on post-crisis reputation. Neither was there an interaction between pitch and speech rate ($F(1, 99) = 2.47$, $p = .12$) (cf. Figure 1).
Second, in the case of a matched crisis response strategy (i.e., rebuild), there were no main effects of pitch \((F(1, 120) = .01, p = .95; M_{low} = 3.13, SD = 1.28 \text{ vs. } M_{high} = 3.05, SD = 1.52)\) and speech rate \((F(1, 120) = .21, p = .65; M_{fast} = 3.15, SD = 1.42 \text{ vs. } M_{slow} = 3.05, SD = 1.38)\) either. However, there was a significant interaction between pitch and speech rate \((F(1, 120) = 9.04, p < .01)\). Two independent samples t-tests allowed looking at this interaction effect of pitch and speech rate in the case of a matched strategy in detail. On the one hand, when communicating with a lower pitched voice, a slow speech rate leads to a better reputation than a fast speech rate \((t(61) = 1.98, p = .05; M_{slow} = 3.41, SD = 1.30 \text{ vs. } M_{fast} = 2.79, SD = 1.18)\). On the other hand, when communicating with a higher pitched voice, a fast speech rate leads to a better reputation than a slow speech rate \((t(59) = -2.25, p < .05; M_{fast} = 3.54, SD = 1.57 \text{ vs. } M_{slow} = 2.69, SD = 1.39)\) (cf. Figure 2). These results answer research questions 1 and 4. Vocal cues do have an impact on post-crisis reputation, but only in the case of a matched (i.e., rebuild) crisis response strategy.
Finally, in order to address the second (Is the impact of vocal cues on post-crisis reputation mediated by the attractiveness of the spokesperson’s voice?) and third research question (Is the impact of vocal cues on post-crisis reputation mediated by empathy?), two mediation analyses were conducted for the three-way interaction between pitch, speech rate and crisis response strategy on post-crisis reputation. To examine these effects, we used PROCESS macro (Hayes, 2013), a modeling tool for SPSS.

First, we examined whether the three-way interaction between pitch, speech rate and crisis response strategy was mediated by attractiveness of the voice. The analysis (5,000 bootstraps, 95% bias-corrected confidence intervals) revealed that the indirect effect of the three-way interaction through attractiveness of the voice was significant ($ab = .78$, $SE = .34$). A 95% bootstrap confidence interval for this indirect effect did not include zero (95% LLCI = .18, 95% ULCI = 1.52), indicating that the interaction effect of crisis response strategy, pitch and speech rate on post-crisis reputation is mediated by attractiveness of the voice. This finding implies that the interplay between verbal and vocal cues affects post-crisis reputation through attractiveness of the CEO’s voice.

Second, we investigated whether the three-way interaction was mediated by the public’s empathy towards the CEO. The analysis (5,000 bootstraps, 95% bias-corrected confidence intervals) revealed that the indirect effect of the three-way interaction through empathy was also significant ($ab = .97$, $SE = .48$). A 95% bootstrap confidence interval for this indirect effect did not include zero (95% LLCI = .06, 95% ULCI = 1.88), indicating that the interaction effect of crisis response strategy, pitch and speech rate on post-crisis reputation is also mediated by
empathy. Therefore, the interplay between verbal and vocal cues affects post-crisis reputation through the public’s empathy towards the CEO.

Discussion

This study explored the impact of vocal cues on post-crisis reputation and examined their interaction with verbal crisis communication. First, contrary to previous findings (e.g., Coombs & Holladay, 1996), the results of this study showed that a matched crisis response strategy (i.e., rebuild in the case of a preventable crisis) did not lead to a significantly better post-crisis reputation than a mismatched crisis response strategy (i.e., deny). Earlier research has shown that using a matched crisis response strategy leads to better attitudes towards the organization when consumers are highly involved with the crisis, but not in the case of low crisis involvement (Claeys & Cauberghe, 2014b). It is possible that participants of this study were not involved enough for the positive effect of a matched strategy compared to a mismatched strategy to occur. Another possibility is that the participants believed that the mismatched, deny strategy was sincere, leading to lower perceptions of crisis responsibility for the organization, and thus leading to a post-crisis reputation that was not significantly worse than in the case of a matched, rebuild strategy. An independent t-test showed that a deny strategy indeed led to a significantly lower attribution of responsibility than a rebuild strategy ($t(225) = -8.30, p < .01; M_{deny} = 4.82, SD = 1.33$ vs. $M_{rebuild} = 6.21, SD = 1.19$). However, another possible explanation is the interplay that seemed to occur between these crisis response strategies and the vocal cues offered by the CEO.

The results indeed showed that there was a significant three-way interaction between pitch, speech rate and crisis response strategy on post-crisis reputation. Further analysis indicated that vocal cues affect post-crisis reputation, depending on the crisis response strategy. Only in the case of a crisis response strategy that matches the crisis type, do vocal cues influence the public’s perception of an organization in crisis.

Two questions arise from these results. First, the question is why vocal cues have an impact on post-crisis reputation only in the case of a matched crisis response strategy (i.e., rebuild). A number of studies have documented that surprising or unexpected information leads to greater processing (e.g., Baker & Petty, 1994). For example, in advertisements, unexpected information leads to higher involvement with the message than expected information (Lee, 2000). Also, unexpected framing of a persuasive message leads to more extensive processing of the message (Smith & Petty, 1996). It is possible that a deny strategy, which is an unexpected choice in the case of a preventable crisis, increases involvement and processing of the message, and thus leads to processing via the central route, which decreases
the impact of peripheral cues, such as pitch and speech rate. However, further research should address this possible explanation.

The second question that arises from the three-way interaction found in this study, is which levels of pitch and speech rate lead to a better post-crisis reputation than others. This study seems to indicate that a high pitch is best combined with a fast speech rate and a low pitch with a slow speech rate. A possible explanation for this result, is that the combinations high-fast and low-slow sound more natural than a high-slow and a low-fast voice. Several studies have shown that the perception of speech rate is influenced by pitch, in that sense that a higher pitch leads to the perception of a faster speech rate (Boltz, 2011; Feldstein & Bond, 1981; Imhof, 2010). As an explanation for this effect, those studies indicate that there is a natural covariation between pitch and speech rate in speech production: slow speech is for example often produced with a lower pitch (Black, 1961). They suggest that by hearing this natural covariation frequently, the perceptual covariation between pitch and speech rate arised. It is thus possible that in crisis communication, the combinations of the vocal cues that do not covary (high-slow and low-fast) sound unnatural, leading to a lower post-crisis reputation. Further research should test this hypothesis.

Finally, this study showed that the three-way interaction between pitch, speech rate and crisis response strategy was mediated by attractiveness of the voice on the one hand and by empathy towards the CEO on the other hand. Earlier research in political communication has already shown that vocal attractiveness affects the perception of political candidates (Surawski & Ossoff, 2006). The results of this study seem to indicate that in the context of crisis communication, vocal attractiveness has an impact on crisis outcome as well and that pitch and speech rate can influence ratings of vocal attractiveness.

Also, empathy seems to play a role in the effect of vocal cues on post-crisis reputation. Vocal cues have an effect on the empathy of the public towards the CEO, which impacts post-crisis reputation. A possible explanation for this effect, is that the voice can convey emotions (van der Meer & Verhoeven, 2014), which could elicit an emotional reaction in the public that leads to empathy towards the CEO (Lawrence et al., 2004). In order to test this suggestion, further research should measure the impact of a spokesperson’s voice in crisis communication on perceived emotions by the public.

Limitations and suggestions for further research

This study has some limitations that further research should address. First, this study had no control conditions with neutral levels of the vocal cues pitch and speech rate. Therefore, the basic effects of separate manipulated cues could not be addressed, because one
manipulated level of a vocal cue was always combined with a manipulated level of the other vocal cue. Further research should include control conditions for vocal cues, in order to check what the effects of manipulated vocal cues are in comparison with neutral, average levels of other vocal cues. Second, pitch and speech rate were investigated in this study because they are two of the most researched cues in impression formation. However, there are other vocal cues that can have an impact (e.g., pitch variability, loudness) (e.g., Apple et al., 1979). Further research should include control conditions for vocal cues, in order to check what the effects of manipulated vocal cues are in comparison with neutral, average levels of other vocal cues. Second, pitch and speech rate were investigated in this study because they are two of the most researched cues in impression formation. However, there are other vocal cues that can have an impact (e.g., pitch variability, loudness) (e.g., Apple et al., 1979). Further research should include control conditions for vocal cues, in order to check what the effects of manipulated vocal cues are in comparison with neutral, average levels of other vocal cues. Second, pitch and speech rate were investigated in this study because they are two of the most researched cues in impression formation. However, there are other vocal cues that can have an impact (e.g., pitch variability, loudness) (e.g., Apple et al., 1979).

Further research should take into account the impact of other vocal cues as well. Third, the crisis scenarios in this study evoked moderate levels of involvement in participants. It is possible that the relative impact of vocal cues and crisis response strategies is different in the case of a low or a high involvement. Prior research indicates that under low involvement, peripheral cues become more important, while under high involvement, the impact of the content of the message increases (e.g., Petty, Cacioppo, & Schumann, 1983). This has already been indicated in social psychological research on vocal cues, where vocal cues affected listeners’ attitudes in the case of low involvement, but not in the case of high involvement (e.g., Gelinas-Chebat & Chebat, 1992). Therefore, further research should test the impact of vocal cues in crisis communication under different levels of crisis involvement. Fourth, this study did not measure perceived emotionality of the speaker and perceived naturalness of the voice. Further research could include these measures, in order to test if those are factors explaining the effects of the combinations of vocal cues on post-crisis reputation.

Conclusions

The results of this study have shown that vocal cues interact with the content of crisis communication and determine post-crisis reputation in times of crisis. Moreover, the impact of the interaction is mediated by attractiveness of the voice and by empathy towards the spokesperson. These results have some important implications for crisis communication research and practice. For research, this study indicates that it is important not only to study the content of crisis communication, but also vocal cues, because they interact with the crisis response strategy and have an impact on post-crisis reputation as well. The interaction between vocal cues and a matched crisis response strategy (i.e., rebuild) is an enlightening finding that is extremely relevant, because literature on crisis communication has emphasized the importance of a crisis response strategy that is matched to the crisis type (e.g., Coombs, 2007), but has neglected its interaction with vocal cues. Moreover, previous research on vocal crisis communication was limited to the impact of one vocal cue, namely voice pitch. This study expands the knowledge on the effect of voice pitch in crisis communication, by investigating the interaction between pitch and speech rate. The results seem to suggest that,
when vocal cues are combined in crisis communication, it is important to combine them in a natural way. Further research should therefore investigate the specific effects of separate vocal cues in order to determine which cues have the most important impact, and which cues should correlate naturally with others. This study thus provides an important contribution to the field of crisis communication research and opens up interesting new opportunities for further research.

For crisis communication practice, the results of this study show that it is important to consider a spokesperson’s vocal characteristics in times of crisis, because they have a significant impact on post-crisis reputation when using a rebuild strategy in preventable crises. Organizations could collaborate with voice coaches in order to train a spokesperson’s voice. In addition, the findings of this study can have very broad applications for corporate communication practice in general. The results are not only relevant for external crisis communication, but can also apply to other situations where spokespersons communicate with stakeholders, such as in the case of internal crisis communication or communication in the context of change management.
References


