Rate of Speech, Intonation, and Pitch:

Investigating the Bias and Cue Effects of Vocal Confidence on Persuasion

Word Count: 9,871

April 12, 2017
Abstract

Across two experiments, we investigated the effects of three vocal qualities on persuasion and the hypothesized underlying psychological mechanisms. Experiment 1 (N = 394) was a 2 (Elaboration: high vs. low) x 2 (Vocal speed: fast vs. slow) x 2 (Vocal intonation: falling vs. rising) between-participants factorial design. Vocal speed and vocal intonation influenced perceptions of speaker confidence as predicted. Under high-elaboration, vocal confidence biased thought-favorability, which then influenced attitudes. Under low-elaboration, vocal confidence directly influenced attitudes as a peripheral cue. Experiment 2 (N = 412) conceptually replicated the bias and cue effects in Experiment 1 using a 2 (Elaboration: high vs. low) x 2 (Vocal pitch: raised vs. lowered) between-participants factorial design. Vocal pitch influenced perceptions of speaker confidence as predicted. These studies demonstrate that changes in three vocal properties influence perceptions of speaker confidence, which then influences attitudes via different mediating processes moderated by amount of thought.

Keywords:
Attitudes, Persuasion, Vocal Confidence, Elaboration Likelihood Model.
For the better part of a century, researchers have investigated variables thought to either facilitate or inhibit attitude change; and more recently, the underlying processes by which these changes take place. While a great deal of research has shown that the content of what we say matters (see, Albarracin & Vargas, 2010; Eagley & Chaiken, 1993; Petty & Wegener, 1998), a number of studies have suggested that how we say something is also important. In fact, researchers have demonstrated that one important feature of oral communication is that our voice allows us to convey a rich variety of information beyond the content of our message (Johnson, Ernde, Scherer, & Klinnert, 1986; Pell, Monetta, Paulmann, & Kotz, 2009). Thus, how we say something should also play a role in how successful we are at persuading others.

Although there are many characteristics of voice that may influence the persuasion process, a growing body of research suggests one characteristic that should play an important role is the extent to which a speaker sounds confident (Brennan & Williams, 1995; Brown, Giles, & Thakerar, 1985; Jiang & Pell, 2014; Kimble & Seidel, 1991; Scherer, London, & Wolf, 1973; Smith & Clark, 1993). Indeed, if one considers that confidence is an important dimension we use to evaluate our own attitudes and thoughts (e.g., Briñol & Petty, 2009; Tormala & Rucker, 2007), it makes sense that confidence may also be an important dimension we use when evaluating other’s communications. For example, when people speak with confidence we may infer they are intelligent, have expertise, and believe what they are saying is valid. Thus, to the extent that confidence is used when making global judgements of others on a variety of attributes, it should also be an important determinant when evaluating whether a persuasive appeal will be successful. But what is it about our voice that communicates confidence?
Vocal Qualities and Speaker Confidence

When investigating which qualities of voice are related to speaker confidence, researchers have asked participants to speak in either a confident or unconfident manner (Scherer et al., 1973), or provided material that impacts self-reported confidence and then observed changes in voice as a result (Kimble & Seidel, 1991). These methodologies have produced converging evidence demonstrating that relative to unconfident speakers, confident speakers typically communicate faster (Brown et al., 1985; Jiang & Pell, 2014; Scherer et al., 1973), louder (Kimble & Seidel, 1991; Scherer et al., 1973), and use falling versus rising intonation at the end of their sentences (Brennan & Williams, 1995).

Vocal Qualities and Persuasion

Although current evidence presents a relatively clear picture of the relationship between specific vocal qualities and perceptions of confidence, much less is known about how these vocal qualities affect persuasion. Of the various qualities of voice linked to confidence, only rate of speech has been examined within the context of persuasion. Within this literature, some studies have found faster speakers are more persuasive (Hausknecht & Moore, 1986; Mehrabian & Williams, 1969; Miller, Maruyama, Beaber, & Valone, 1976; Moore et al., 1986; Smith & Shaffer, 1991; 1995), whereas others have found no difference between speakers talking at a rapid versus normal rate of speed (Gunderson & Hopper, 1976; Woodall & Burgoon, 1983).

Unfortunately, the techniques often used to manipulate rate of speech present a variety of methodological issues that complicate interpretations of the data. For example, researchers have instructed the speaker to talk fast or slow (e.g., Miller et al., 1976, Woodall & Burgoon, 1983), used audio-visual stimuli that presented multiple variables designed to either singularly enhance or reduce persuasion (e.g., Gunderson & Hopper, 1976; Woodall & Burgoon, 1983), or forcibly
increased speech rate by compressing an analog audio file through re-recording the original track (e.g., Hausknecht & Moore, 1986; Moore et al., 1986; Smith & Shaffer, 1991; 1995). These methodologies are problematic because when instructed to speak either fast or slow, a speaker may inadvertently change multiple properties of voice (e.g., volume, intensity, pitch) and therefore confound the manipulation. Thus, whether persuasion is driven by changes in one or multiple vocal properties becomes impossible to determine. Second, using audio-visual stimuli that presents multiple variables designed either to singularly enhance or reduce persuasion is problematic because it obscures the influence of any one variable on persuasion. Finally, it is not clear whether forcibly compressing an audio file altered certain parameters of voice that may have influenced persuasion.

Curiously, although the vocal perception literature suggests that rate of speech is an important determinant of confidence, and the attitudes literature indicates confidence is an important attribute people use when making decisions, research has not examined whether rate of speech influences persuasion via its impact on perceptions of speaker confidence. Moreover, no research has investigated whether combining different vocal qualities influences perceptions of confidence in an additive or interactive fashion. Finally, only a small handful of studies have experimentally tested the mechanisms governing the relationship between rate of speech and persuasion. Some researchers have suggested rate of speech enhances persuasion by serving as a cue for enhanced credibility (Miller et al., 1976; Smith & Shaffer, 1995), whereas others have proposed amount of processing as the likely mechanism (Hausknecht & Moore, 1986; Moore et al., 1986; Smith & Shaffer, 1991). Thus, there is little agreement regarding when and why rate of speech – and other qualities of voice tied to perceptions of confidence, influence persuasion.
Vocal Confidence and Persuasion: A Theoretical Framework

One theory that may be useful in organizing these disparate findings is the Elaboration Likelihood Model (ELM; Petty & Cacioppo, 1986; Petty & Wegener, 1998). The ELM postulates that a variety of different processes can guide attitude change. Which specific process emerges is a function of where an individual falls on the elaboration continuum, which is determined by an individual’s ability and motivation to effortfully process information. When both ability and motivation are high, this typically leads a person to carefully scrutinize the merits of an argument. In contrast, when ability and/or motivation are low, careful scrutiny of an argument’s merits is less likely.

Applying this to rate of speech and other qualities of voice tied to confidence, when a person is able and motivated to carefully consider the merits of an issue, the ELM suggests they are more likely to consider both the relevance and quality of all information that comes to mind as it relates to the persuasive communication. Under these conditions the ELM proposes that a variable can influence persuasion via biasing the recipient’s thoughts to be more or less favorable toward the message.¹ Take, for example, the topic of nuclear power. The confidence with which the speaker delivers the message in no way directly informs the listener whether nuclear power is safe, economical, environmentally friendly, and so forth. Nonetheless, how the speaker communicates the arguments could influence the listener to process arguments in either a more or less favorable manner. For example, the listener may infer that a confident speaker is exceptionally knowledgeable or passionate and thus reason the speaker may be sharing valuable information. Consequently, the listener may focus on the strengths of the argument and thus approach the message in a very confirmatory way. This may positively bias thoughts towards the message, and these positive thoughts will then be used as a basis for the attitude. However, if the
speaker sounds unconfident, the listener may recognize this and reason that the speaker may not be providing very accurate information. Consequently, the listener may focus on the weaknesses in the arguments, which then negatively biases their thoughts towards the message. Research suggests the biasing effects of a variable should be most effective when a message is composed of moderately strong arguments (Chaiken & Maheswaran, 1994). Importantly, research has not investigated whether rate of speech – or other vocal properties associated with confidence – may impact persuasion because of how it influences our perceptions of speaker confidence, which in turn can bias the favorability of message-relevant thoughts.

In contrast, when a person is unable and/or unmotivated to think carefully, the ELM suggests they are more likely to attend to simple cues in the environment to guide their response to the message. Applying this to rate of speech and other qualities of voice tied to confidence, a listener may evaluate whether a speaker sounds confident and then use this perception as a peripheral cue to directly infer their attitude toward the advocated position in the absence of any careful evaluation of the quality of the advocacy. Thus, our expectation is that rate of speech – and other vocal properties associated with confidence, should elicit either favorable or unfavorable attitudes without directly impacting the favorability of the listener’s thoughts.

**Overview of the Present Research**

At a broad level, this research examined two mechanisms suggested by the ELM that describe how rate of speech, as well as other vocal properties tied to confidence, can influence our perceptions of speaker confidence, which in turn affect persuasion. Several goals guided our research: First, we sought to investigate whether vocal properties found to influence perceptions of confidence beyond rate of speech, such as vocal intonation and vocal pitch, might also influence persuasion. Importantly, we wanted to determine if these vocal properties influenced
persuasion because of how they affected perceptions of speaker confidence, and further, whether the mechanisms by which confidence influenced persuasion differed based on the recipient’s ability and motivation to process the message. In addition, we wanted to test whether different combinations of two vocal properties influenced perceptions of confidence and persuasion in an additive or interactive fashion. Finally, we sought to address methodological issues with vocal manipulations in prior research by using a digital recording and editing process that enabled greater precision when manipulating specific qualities of voice.

Experiment 1

Experiment 1 investigated how rate of speech and vocal intonation influenced persuasion using a message comprised of moderately strong arguments. More specifically, we tested the processes by which these hallmarks of vocal confidence influenced persuasion under high and low thought based on the predictions made by the ELM. Ratings of speaker confidence and cognitive responses were tested as mediators of this process.

Method

Participants and Design

Three hundred ninety-four undergraduates were recruited in exchange for course credit. Participants were randomly assigned to conditions in a 2 (Elaboration: high vs. low) x 2 (Vocal speed: increased vs. decreased) x 2 (Vocal intonation: falling vs. rising) between-participants factorial design.

Procedure

Participants listened to an audio passage describing a policy that allowed students to reduce their tuition in exchange for working as part-time university staff members. The passage was modelled after pro-attitudinal arguments on the same topic that prior research had shown
were perceived as comparatively strong (Baker & Petty, 1994; Clark, Wegener, & Fabrigar, 2008). We modified the arguments to be moderately strong because prior research suggests the biasing effect of a variable (e.g., vocal confidence) are most effective when arguments are only moderately strong (Chaiken & Maheswaran, 1994).

An important goal of Experiment 1 was to investigate how vocal hallmarks of confidence affect persuasion at the end points of the elaboration continuum. Thus, when creating our high-elaboration condition, we sought to maximize processing ability by placing participants in semi-private conditions, thereby minimizing audio and visual distractions to the greatest extent possible. Motivation to process the arguments was maximized by framing the message as high in personal relevance. Specifically, we informed participants their university was considering the policy in question for implementation the following year (Petty & Cacioppo, 1979; Petty & Cacioppo, 1984). In contrast, our low-elaboration condition sought to minimize processing ability using a distraction task that required memorizing an 8-digit number and recalling it at the end of the experiment (e.g., Gilbert & Osborne, 1989). Motivation was minimized by framing the message as low in personal relevance. Participants were told that although several universities had adopted the program, their university was not considering this program.

The audio passage was delivered by a female speaker recruited from the university drama program. A digital recording and editing program (PRAAT©) was used to create the audio passage, thus enabling us to manipulate only vocal speed and intonation without affecting other vocal qualities. When recording the passage, the speaker talked at her normal rate of speech and delivered the content as naturally as possible. Vocal speed was manipulated by increasing speech rate by 10% (191 WPM) and decreasing speech rate by 15% (149 WPM) relative to baseline in the original audio recording. Vocal intonation was manipulated by raising or lowering the
intonation in the speaker’s voice on the last word in 10 target sentences. Prior research conducted in our lab indicated these manipulations were successful. Fully crossing these manipulations produced four audio recordings in which the speaker either spoke fast with rising or falling intonation, or slow with rising or falling intonation. Following the audio passage, participants assigned to low-elaboration were asked to enter the number they were given. Next, all participants answered three filler questions regarding the speaker’s stylistic delivery of the message. Participants then evaluated the speaker on two attributes and two qualities of voice. Attitude toward the proposed tuition-reduction plan was then measured. Finally, participants listed up to 12 thoughts.

Measures

Stylistic Qualities.

Three questions assessing the stylistic qualities of the speaker were included to mask the intent of the study. Question order was randomized. Participants evaluated the speaker’s clarity of presentation, complexity of vocabulary, and organization of points. Responses were made on a 1 to 7 scale, where 1 represented Very unclear/basic/disorganized, and 7 represented Very clear/complex/organized, respectively.

Speaker Attributes and Vocal Qualities.

Four questions were presented, of which two (i.e., vocal confidence, rate of speech) were of interest and two (i.e., intelligence, loudness) were filler items. Question order was randomized. Responses were made on a 1 to 7 scale, where 1 represented Very unconfident/slow/quiet/unintelligent, and 7 represented Very confident/fast/loud/intelligent, respectively.
**Attitude Scale.**

Attitudes were measured using an 8-item scale consisting of different words reflecting general and undifferentiated positive or negative evaluation (Crites, Fabrigar, & Petty, 1994). Half of the words implied positive evaluations (e.g., good, positive), whereas half implied negative evaluations (e.g., dislike, undesirable). Final scores were created by reverse coding the negative items, then averaging the scores across all scale items. Responses were recorded on a scale ranging from 1 = *Not at all*, to 7 = *Definitely*. Cronbach’s α for the attitude scale was .93.

**Thought Listing and Rating Task.**

A thought-listing and rating task (e.g., Cacioppo & Petty, 1981) asked participants to list up to twelve thoughts that came to mind while listening to the audio passage and then rate the favorability of each thought as *Positive*, *Negative*, or *Neutral*, as it applied to the university service plan. Two independent raters also coded participant’s thoughts on favorability and relevance to the topic, while blind to condition (e.g., Cacioppo & Petty, 1979; Petty & Cacioppo, 1984). Interrater agreement on both thought-favorability (92%) and thought relevance (88%) was high. Final values were calculated by averaging the scores for both independent raters on each dimension (e.g., Petty, Cacioppo, & Heesacker, 1981). Based on the coding assigned by the independent raters, an index of thought favorability was computed using the following formula:

\[
\text{Thought Favorability} = \frac{\text{Number of positive relevant thoughts} - \text{Number of negative relevant thoughts}}{\text{Total number of thoughts}}
\]

Using scores from the thought-rating task, we evaluated the success of our elaboration manipulation by examining the *proportion* of topic-relevant thoughts compared with all thoughts generated by the participant. \(^4\)
Results

Elaboration Manipulation Check

Our first goal was to evaluate the success of our elaboration manipulation. If successful, participants carefully processing the message should have a significantly greater proportion of topic-relevant thoughts than participants not carefully processing the message. This was tested by conducting an ANOVA with vocal speed, intonation, and elaboration as the independent variables and the proportion of topic-relevant thoughts as the dependent variable. As expected, increased processing led to a greater proportion of topic-relevant thoughts by participants assigned to high-elaboration ($M = .82, SE = .02, 95\% CI .79, .86$), than low-elaboration ($M = .75, SE = .02, 95\% CI .71, .78$), $F(1, 380) = 7.88, p < .01$, partial $\eta^2 = .02$. Although we had no basis to expect further effects, a greater proportion of topic-relevant thoughts emerged in response to falling ($M = .85, SE = .02, 95\% CI .81, .89$), versus rising intonation ($M = .72, SE = .02, 95\% CI .68, .76$), $F(1, 380) = 19.71, p < .001$, partial $\eta^2 = .05$. No further effects reached significance. These data provide evidence confirming the success of our elaboration manipulation.

Evaluation of Vocal Confidence Measure

Next, because prior research has shown that changes in rate of speech and vocal intonation influence perceptions of speaker confidence, we wanted to confirm that both vocal manipulations replicated these effects. This was tested by conducting an ANOVA with vocal speed, intonation, and elaboration as the independent variables and ratings of speaker confidence as the dependent variable. Two critical effects were expected. First, a main effect of vocal speed was predicted. Confirming expectations, the speaker with an increased rate of speech ($M = 5.39, SE = .11, 95\% CI 5.18, 5.59$), was rated as more confident than the speaker with a decreased rate of speech ($M = 4.51, SE = .11, 95\% CI 4.30, 4.72$), $F(1, 394) = 34.26, p < .001$, partial $\eta^2 = .08$. 
Second, we predicted a main effect of vocal intonation. As anticipated, the speaker with falling intonation ($M = 5.51, SE = .11, 95\% \text{ CI} 5.30, 5.72$), was rated as more confident than the speaker with rising intonation ($M = 4.39, SE = .11, 95\% \text{ CI} 4.20, 4.59$), $F(1, 386) = 55.75, p < .001, \text{ partial } \eta^2 = .13$. No main effect of elaboration or interaction between vocal qualities and elaboration was expected and none emerged. The absence of an interaction between elaboration and vocal qualities is important because it suggests that perceptions of speaker confidence were not influenced by the amount of processing. Interestingly, no interaction emerged between vocal speed and vocal intonation, $F(1, 386) = .22, p = .64, \text{ partial } \eta^2 < .01$, thus suggesting the impact of these variables on perceptions of speaker confidence was additive rather than interactive. Finally, a three-way interaction was not expected and did not emerge.

*The Effects of Vocal Qualities and Elaboration on Attitudes*

Our next step was to investigate the effects of vocal qualities and elaboration on participant’s attitudes toward the university service plan. This was tested by conducting an ANOVA with vocal speed, intonation, and elaboration as our independent variables and attitude as the dependent variable. Although prior research suggests the possibility that several effects could emerge, it is important to consider that because attitudes assume the most distal position in our assumed causal model, to the extent that any effects of vocal qualities do emerge, these effects should be comparatively weak. We hypothesized that persuasion may be greater in response to a speaker who communicated at a comparatively rapid pace and/or used falling intonation at the end of a sentence. However, no main effect of vocal speed, $F(1, 386) = .07, p = .79, \text{ partial } \eta^2 = .00$, or vocal intonation, $F(1, 386) = 1.34, p = .25, \text{ partial } \eta^2 = .00$ was found. No main effect of elaboration or interaction between vocal qualities and elaboration was expected because vocal qualities were hypothesized to have similar effects on attitudes but
through different mediating processes at each level of elaboration. As anticipated, none of these effects reached significance. Taken together, the absence of a main effect for either vocal quality was not surprising. Critical to our model, however, was the expectation that neither vocal quality would interact with elaboration, which is exactly what we found.

_Vocal Confidence as a Biasing Factor and Peripheral Cue_

In order to test the most central prediction of our theoretical framework, we sought to evaluate whether the process by which vocal speed and vocal intonation influenced persuasion differed under high – and low-elaboration. This was tested by conducting a multi-sample structural equation model using Lisrel 9.20 (Joreskog & Sorbom, 2014). First, participants were divided into high and low-elaboration conditions. Next, the covariance matrix for each level of elaboration was computed using the measured variables designated as exogenous (i.e., vocal speed and vocal intonation) and endogenous (i.e., ratings of speaker confidence, cognitive responses, and attitude). When conducting our analyses, the model depicted in Figure 1 was fit simultaneously to both groups. Using Maximum Likelihood to estimate the model parameters, we tested whether various paths were significantly different from one another across levels of elaboration by placing equality constraints on specific coefficients of interest. A chi-square difference test was used to evaluate the fit of our constrained model against its unconstrained counterpart. A significant chi-square statistic indicates the particular coefficients being compared were significantly different. Finally, recall that the earlier ANOVA investigating the joint effects of vocal speed and vocal intonation on perceptions of confidence revealed that these variables influenced confidence in an additive rather than interactive fashion. Thus, these models assume additive effects of voice. The results for both path models are represented in Figure 2 and report unstandardized coefficients for each path. Dotted lines indicate non-significant paths.
According to the Elaboration Likelihood Model, the process by which a variable influences persuasion differs based on the extent of careful thought regarding the merits of an argument. Under high-elaboration, the ELM suggests a variable should bias thought-favorability, which in turn is used as a guide when forming subsequent attitudes. Thus, our prediction was that vocal speed and vocal intonation would influence perceptions of speaker confidence, which should then bias the recipient’s thoughts and in turn influence their attitude. Importantly, recall that the ELM suggests a different process should occur under low-elaboration. In this case, although vocal speed and vocal intonation should still influence perceptions of speaker confidence, the difference here is that confidence no longer biases thought-favorability but rather directly impacts the recipient’s attitude as a peripheral cue.

We start by investigating the pattern of effects the ELM suggests should emerge under high-elaboration. As found in panel A of figure 2, both vocal speed, $b = .83, SE = .21, p < .001$, and vocal intonation, $b = .92, SE = .21, p < .001$, influenced perceptions of speaker confidence as predicted. Next, a confident sounding speaker should bias thoughts to be more positive toward the message. As expected, this is exactly what we find, $b = .15, SE = .03, p < .001$. Finally, because participants were presumed to be engaged in careful thought, we would expect thoughts to have a robust effect on attitudes, which the data supported, $b = 1.29, SE = .11, p < .001$. The final path critical to evaluating our perspective is the direct effect of speaker confidence on attitude. Because this is a high-elaboration situation, the ELM predicts that speaker confidence should bias thought-favorability but have little direct impact on attitude as a peripheral cue. As anticipated, this cue effect was non-significant, $b = .02, SE = .05, p = .73$. In line with our theoretical framework, no further paths reached significance. This is important because it suggests that rate of speech and intonation do not directly impact attitude but rather do so
indirectly because of how they influenced perceptions of speaker confidence. Taken together, these data reveal a pattern of effects that support our theoretical framework.

We now examine the pattern of effects that emerged under low-elaboration. As found in panel B of Figure 2, once again vocal speed, \( b = .89, SE = .23, p < .001 \), and vocal intonation, \( b = 1.31, SE = .23, p < .001 \), influenced perceptions of speaker confidence as expected. According to our theoretical framework, people should be equally adept at detecting changes in speech rate and intonation regardless of how carefully they are thinking. Thus, we expected both variables should have similar effects on confidence across levels of elaboration. Indeed, comparing these coefficients with their counterparts in panel A revealed similar numerical values. Testing these coefficients across levels of elaboration indicated the effects of vocal speed, \( \chi^2 = .03 \) \((df = 1, N = 371, p = .85)\), and vocal intonation, \( \chi^2 = 1.67 \) \((df = 1, N = 371, p = .20)\), were of comparable magnitude. This suggests that amount of processing did not differentially impact participant’s ability to detect changes in the speaker’s voice.

Next, when ability and/or motivation to carefully process a message are lacking, speaker confidence should have a comparatively weak effect on thought-favorability. Indeed, comparing this coefficient with its counterpart in panel A suggests this was the case, \( b = -.03, SE = .03, p = .38 \). Testing these coefficients across levels of elaboration confirmed that confidence was a significantly better predictor of thoughts under high – compared with low-elaboration, \( \chi^2 = 15.85 \) \((df = 1, N = 371, p < .001)\). Importantly, although thoughts may still have some influence on attitudes under low-elaboration, the ELM suggests this effect should be weaker than under high-elaboration. As anticipated, a smaller coefficient emerged relative to its counterpart in panel A, \( (b = .79, SE = .10, p < .001) \). Once again, testing these coefficients across elaboration revealed the expected effect, \( \chi^2 = 11.10 \) \((df = 1, N = 371, p < .001)\).
Finally, under low-elaboration, speaker confidence should directly influence attitudes as a peripheral cue. Thus, here we would expect a significant positive coefficient, which emerged as anticipated, $b = .16, SE = .05, p < .01$. When comparing this coefficient with its counterpart in panel A, tests confirmed this effect was significantly more powerful under low – than high-elaboration, $\chi^2 = 4.46 (df = 1, N = 371, p = .03)$. No further paths reached significance. These data support our theoretical framework by showing that under low-elaboration, confidence does not influence thought-favorability but rather directly affects attitude as a peripheral cue.

Discussion

These data replicate past research (Brennan & Williams, 1995; Jiang & Pell, 2014; Scherer et al., 1973), by demonstrating that rate of speech and vocal intonation are important determinants of speaker confidence. Importantly, this experiment is the first to demonstrate that these hallmarks of vocal confidence influence persuasion because of how they affect the recipient’s perception of speaker confidence. Using digital technology to create more precise manipulations of voice, these data provide the first evidence that different combinations of two vocal properties influence confidence and persuasion in an additive rather than interactive fashion. Notably, as the ELM predicts, these data show that when effortfully processing a message, perceptions of speaker confidence can bias thought-favorability. However, when processing is reduced, speaker confidence does not bias thought-favorability but rather directly affects attitudes as a peripheral cue.

Experiment 2

Given the multifaceted nature of voice, it seems plausible that additional qualities beyond rate of speech and intonation may affect perceptions of speaker confidence and in turn influence persuasion in a similar manner as shown in Experiment 1. For example, research has found that
changes in vocal pitch reliably influence listener’s judgements of a speaker on various dimensions, including competence (Brown, Strong, & Rencher, 1973), honesty (Streeter, Krauss, Geller, Olson, & Apple, 1977), and anxiety (Apple, Streeter, & Krauss, 1979), such that raised pitch elicited more negative evaluations on each dimension. Given that raised pitch is associated with negative evaluations on both competence and anxiety, and anxiety and confidence are inversely related, it follows that listeners may associate increased confidence with lowered pitch. Of course, if pitch is also a determinant of confidence, it makes sense that pitch should influence persuasion in a similar manner as vocal speed and vocal intonation when message processing occurs at either end of the elaboration continuum. Thus, an important goal of Experiment 2 was to show that the same psychological processes that emerged in Experiment 1 could also be extended to vocal qualities not previously investigated in the vocal confidence literature. Finally, Experiment 2 sought to provide a second demonstration of the multiple roles (i.e., bias, peripheral cue) by which speaker confidence can affect persuasion under high and low-elaboration through a conceptual replication.

Method

Participants and Design

Four hundred twelve undergraduates were recruited in exchange for course credit and randomly assigned to condition in a 2(Elaboration: high vs. low) x 2(Vocal pitch: raised pitch vs. lowered pitch) between participants factorial design.

Procedure

Participants listened to an audio passage that discussed the benefits of using phosphate-based laundry detergent (Shavitt & Brock, 1986). Similar procedures were used to manipulate processing ability as in Experiment 1. Motivation to process the message was influenced by
using a manipulation of personal responsibility (e.g., Petty, Harkins, & Williams, 1980). Under high-elaboration, motivation was enhanced by informing participants that because very few students would be completing the survey, they may be one of the only students offering feedback. Thus, their feedback was especially important to the researchers. By contrast, under low-elaboration, motivation was reduced by informing participants that because so many students would be completing this survey, their responses may be discarded. Thus, any information they provide may not be read.

The audio passage was delivered by a male speaker recruited from the university psychology department. The same program (PRAAT©) was used to create the audio passage as in Experiment 1, thus enabling manipulation of the speaker’s pitch without altering other vocal qualities. The speaker spoke naturally and at his normal rate of speed. Vocal pitch was manipulated by raising the pitch in the speaker’s voice by 120 hertz or lowering the pitch by 20 hertz relative to the speaker’s natural baseline. Research conducted in our lab using the same speaker and manipulations of pitch produced the expected effects on ratings of confidence. Following the audio passage, participants assigned to low-elaboration entered the number they were given. Next, attitude toward phosphate-based laundry detergents was measured. Participants then evaluated different attributes of the speaker as well as the speaker’s voice. Finally, participants listed up to 10 thoughts, then rated the favorability of those thoughts as either positive, negative, neutral, or unrelated (Briñol, Petty, & Tormala, 2004).

Measures

Attitude Scale.

The attitude scale was identical to that used in Experiment 1. Cronbach’s α for the attitude scale was .92.
**Speaker Attributes and Vocal Qualities Questionnaire.**

Of twelve items, three (i.e., confidence, pitch, naturalness)\(^8\) were of primary theoretical interest. Nine items (i.e., age, honesty, sincerity, intelligence, knowledgeable, competence, trustworthiness, credibility, and anxiety) were included for exploratory purposes. Ratings of pitch, naturalness, and confidence were always presented first. All remaining questions were presented in random order. Ratings of speaker pitch used a scale ranging from \(-3 = \text{Very low}\), to \(+3 = \text{Very high}\). All remaining questions used identical scaling (i.e., \(1 = \text{Not at all}\), to \(7 = \text{A great deal}\)), except ratings of speaker age, which used an open-ended format.

**Thought Listing and Rating Task.**

The thought-listing and rating tasks were identical to those used in Experiment 1, but allowed participants to list a maximum of 10 rather than 12 thoughts. Identical coding procedures as Experiment 1 were used for all indices. Evaluations of thought-favorability and thought relevance used the same independent raters and identical procedures as Experiment 1. Interrater agreement on both thought-favorability (89\%) and thought relevance (87\%) was high.

**Results**

**Elaboration Manipulation Check**

As in Experiment 1, we first sought to evaluate the success of our elaboration manipulation. Once again, carefully processing the message should result in a significantly greater proportion of relevant thoughts than not carefully processing the message. This was tested by conducting an ANOVA with vocal pitch and elaboration as the independent variables and the proportion of relevant thoughts as the dependent variable. As expected, participants assigned to high-elaboration \((M = .71, SE = .02, 95\% \text{ CI } .67, .76)\), generated a greater proportion of relevant thoughts than low-elaboration \((M = .62, SE = .02, 95\% \text{ CI } .57, .67)\), \(F(1, 401) = 7.83,\)
Against expectations, a main effect of vocal pitch emerged, $F(1, 401) = 14.35, p < .001$, partial $\eta^2 = .04$, revealing that lowered pitch ($M = .73, SE = .02, 95\% CI .68, .78$), elicited a greater proportion of relevant thoughts than raised pitch, ($M = .62, SE = .02, 95\% CI .56, .65$). No vocal pitch by elaboration interaction was found, $F(1, 401) = .38, p = .54$, partial $\eta^2 = .00$. Similar to Experiment 1, these data provide evidence that our elaboration manipulation was successful.

**Evaluation of Vocal Confidence Measure**

Recall that prior research has found an association between raised pitch and perceptions of greater anxiety. Given that confidence and anxiety are inversely related, we hypothesized that lowered pitch should elicit perceptions of higher speaker confidence. This was tested by conducting an ANOVA with vocal pitch and elaboration as the independent variables and ratings of speaker confidence as the dependent variable. Confirming expectations, lowered pitch ($M = 4.63, SE = .10, 95\% CI 4.43, 4.83$), elicited higher ratings of confidence than raised pitch ($M = 3.12, SE = .10, 95\% CI 2.92, 3.32$), $F(1, 408) = 109.43, p < .001$, partial $\eta^2 = .21$. Although we had no basis to expect a main effect of elaboration, nonetheless, participants assigned to low-elaboration ($M = 4.02, SE = .10, 95\% CI 3.82, 4.22$), provided higher ratings of confidence than high-elaboration ($M = 3.73, SE = .10, 95\% CI 3.53, 3.92$), $F(1, 408) = 4.17, p = .04$, partial $\eta^2 = .01$, Importantly, no interaction between vocal pitch and elaboration emerged, $F(1, 408) = .001, p = .97$, partial $\eta^2 = .00$. This suggests that the effects of vocal pitch on ratings of confidence was equal across levels of elaboration.

**The Effects of Vocal Pitch and Elaboration on Attitudes**

Next, we examined the effects of vocal pitch and elaboration on participant’s attitudes by conducting an ANOVA with vocal pitch and elaboration as our independent variables and
attitude as the dependent variable. Recall that in Experiment 1 we hypothesized the possibility of an overall effect of speech rate and intonation on attitude with the caveat that their distal position in the causal chain may preclude this effect from emerging. Similar logic guided our predictions with vocal pitch. Results indicated lowered pitch ($M = 4.89, SE = .08, 95\% CI 4.74, 5.04$), elicited more persuasion than raised pitch ($M = 4.60, SE = .08, 95\% CI 4.45, 4.75$), $F(1, 408) = 7.26, p < .01, \text{partial } \eta^2 = .02$, thus suggesting the effects of vocal pitch were sufficiently robust to produce differences in attitude. No further effects were predicted and none emerged.

**Vocal Confidence as a Biasing Factor and Peripheral Cue**

As in Experiment 1, our critical analysis utilized a multi-sample structural equation model that tested the psychological processes our theoretical framework predicts should emerge under different amounts of thought. We anticipated similar patterns as those in Experiment 1. The results for both path models are represented in Figure 3 and report unstandardized coefficients for each path. Dotted lines indicate non-significant paths.

Under high-elaboration, as seen in panel A of figure 3, vocal pitch, $b = 1.52, SE = .22, p < .001$, influenced perceptions of speaker confidence as predicted. Next, confidence should bias the favorability of a person’s thoughts. As anticipated, this is exactly what we found, $b = .16, SE = .03, p < .001$. Importantly, because participants were presumed to be engaged in careful thought, we would expect thoughts to have a robust effect on post-message attitude. Indeed, the data revealed this was the case, $b = .88, SE = .11, p < .001$. Finally, recall that under high-elaboration, speaker confidence should bias thought-favorability but have little direct impact on attitude as a peripheral cue. Indeed, this cue effect did not emerge, $b = .01, SE = .05, p = .82$. In line with our theoretical framework, no further paths reached significance. As in Experiment 1, these data reveal a pattern of effects that support our theoretical framework.
Turning our attention to low-elaboration, as found in panel B of Figure 3, vocal pitch influenced perceptions of speaker confidence as predicted, $b = 1.46, SE = .21, p < .001$. Because our theoretical framework suggests that people should be equally adept at detecting changes in pitch regardless of how carefully they are thinking, we expected similar effects of pitch on confidence across levels of elaboration. Indeed, comparing this coefficient with its counterpart in panel A revealed similar numerical values, $\chi^2 = .04$ ($df = 1, N = 258, p = .84$).

Next, recall that when ability and/or motivation to carefully process a message are lacking, our theory predicts speaker confidence should have a comparatively weak effect on thought-favorability, $b = .05, SE = .04, p = .24$. Comparing these coefficients across levels of elaboration confirmed that confidence was a better predictor of thoughts under high – compared with low-elaboration, $\chi^2 = 5.20$ ($df = 1, N = 258, p = .02$). Importantly, although thoughts should still have some influence on attitudes under low-elaboration, $b = .37, SE = .10, p < .001$, this effect should be weaker than under high-elaboration. Once again, testing these coefficients across elaboration revealed the expected effect, $\chi^2 = 10.81$ ($df = 1, N = 258, p < .001$).

Finally, under low-elaboration, speaker confidence should directly influence attitudes as a peripheral cue. Thus, we would expect a significant positive coefficient, which emerged as anticipated, $b = .17, SE = .05, p < .01$. Importantly, tests confirmed this effect was more powerful under low – than high-elaboration, $\chi^2 = 4.59$ ($df = 1, N = 258, p = .03$). No further paths reached significance. Importantly, these data replicate the patterns in Experiment 1 while also extending the findings to an additional vocal determinant of perceived speaker confidence.

**Testing Potential Alternative Mediators**

Despite providing evidence that perceptions of speaker confidence mediate the effects of speech rate, intonation, and pitch on persuasion, this does not preclude the possibility that other
perceived attributes of the speaker could also mediate this relationship. Indeed, the possibility exists that confidence may be correlated with other perceived attributes of the speaker and thus only appear as if it functions in a mediating role because it serves as a proxy for a different attribute. In order to test this possibility, Experiment 2 measured a variety of other attributes to bolster our claim that confidence serves as the true mediator driving the effects of these vocal qualities on persuasion. Specifically, participants’ attributes of the speaker were also measured along the dimensions of credibility, competence, intelligence, knowledgeability, trustworthiness, anxiety, age, honesty, and sincerity.

To evaluate the possibility that one or more of these additional attributes may serve in a mediating role and produce the same bias and cue effects predicted by our model under conditions of high and low-elaboration, we conducted an identical analyses to the multi-sample structural equation model reported in Figure 2 for each of the aforementioned attributes. The results of these analyses revealed that none of the aforementioned attributes produced the expected pattern of effects under either high – or low-elaboration, with the exception of sincerity. More specifically, in each analysis, at least one but in many cases several of the expected paths under high – and low-elaboration failed to reach significance. However, in the case of sincerity, the data did in fact reveal the anticipated mediational patterns via the bias and cue effects predicted by our model under high – and low-elaboration. Thus, to evaluate whether confidence may serve as a proxy for sincerity and thus not be the actual mediator driving the effects of vocal pitch on persuasion, we conducted the identical analyses but in this case included both confidence and sincerity as potential mediators. Supporting our perspective that perceptions of speaker confidence are responsible for the effects of vocal qualities on persuasion, the data indicated that the bias and cue effects for confidence predicted by our model under high – and
low-elaboration remained significant, while at least one path under each level of elaboration for sincerity dropped to non-significance. This suggests that confidence has a robust effect on persuasion under both bias and cue roles even when controlling for a potential mediator that showed patterns of effects consistent with our theoretical model.

Discussion

These data present the first evidence that changes in vocal pitch influence perceptions of speaker confidence. Moreover, our data suggests that different vocal qualities that influence perceptions of confidence also affect persuasion by the same bias and cue processes the ELM suggests should emerge based on the extent of careful thought. Critically, by testing potential alternative mediators, these data strengthen our claim that confidence is a mediator responsible for the effects of pitch on persuasion.

General Discussion

Summary of Findings

These data replicated past research (Brennan & Williams, 1995; Brown et al., 1985; Smith & Clark, 1993; Jiang & Pell, 2014; Scherer et al., 1973), by providing clear evidence that changes in rate of speech and vocal intonation impact perceptions of speaker confidence. Moreover, we demonstrated that perceptions of confidence are also influenced by a property of voice not previously associated with confidence: vocal pitch. These data also provide some evidence that, when combined, different vocal properties work together in an additive rather than interactive fashion to influence perceptions of speaker confidence. More broadly, these data support the predictions derived from the ELM regarding the processes by which vocal confidence should affect persuasion. Specifically, when carefully evaluating the merits of an argument, confidence biases thought-favorability but does not influence attitude as a peripheral cue. By contrast, when
not carefully evaluating an argument, confidence does not influence thought-favorability but rather directly influence attitudes as a peripheral cue. Across two experiments, these bias and cue effects were confirmed in the context of speech rate, intonation, and pitch.

Theoretical and Practical Implications

Oral exchanges constitute a large portion of our daily communication. Both intuitively and empirically, we know the impact of oral communication goes beyond the content of what we say. Indeed, research has shown that vocal characteristics such as tone and emotionality also play an important role in oral exchanges (Johnson et al., 1986; Mandal, 2008; Pell et al, 2009). Unfortunately, despite the large body of literature examining variables that influence persuasion, researchers have largely neglected to investigate how qualities of voice affect persuasive communications. The current research bridges some gaps in this literature by demonstrating the multifaceted relationship between vocal characteristics that reflect confidence and their effects on persuasion. These data address this relationship in a number of important ways.

First, although the persuasion literature has shown that rate of speech can have an effect on attitudes, little work has been done to identify why this vocal property should matter. Drawing upon research from the vocal perception literature, these studies argued that speaker confidence is the critical mechanism by which rate of speech has its effect on attitudes. Moreover, our data provide evidence that not only does rate of speech have its effect on attitudes because of how it influences perceptions of speaker confidence, but that this relationship can also be extended to other properties of voice that serve as determinants of confidence, such as intonation and pitch.

Importantly, recall that prior research attempted to identify a single process to explain the effects of rate of speech on persuasion. Our data, however, present evidence demonstrating that
this way of conceptualizing the relationship between rate of speech, as well as other properties of voice tied to confidence, and persuasion is too simplistic because it does not capture the different psychological processes that emerge under different levels of thought. For example, when carefully thinking, rate of speech, intonation, and pitch influence the listener’s perceptions of speaker confidence, which then biases the favorability of their message-relevant thoughts. In turn, thoughts inform the listener’s attitude toward the target. By contrast, when not carefully thinking, vocal properties still influence perceptions of speaker confidence, which then directly influences the listener’s attitude but does not bias thought-favorability. Thus, our data revealed that although rate of speech, intonation, and pitch each have a similar influence on confidence when listeners are either carefully processing a message or they are not, the mechanism by which confidence effects attitudes differs based on the amount of processing.

Finally, at a more general level, these experiments are another test of the ELM’s ability to organize a broad range of variables into a cohesive structure that predicts how a variable will function under a diverse set of conditions and the underlying psychological processes responsible for its effect on attitudes. This notion of multiple roles has been applied to a host of variables across myriad topics with great success (see, Petty, Rucker, Bizer, & Cacioppo, 2004). Our data show that this framework also applies in the context of vocal properties, thus providing further support for the utility of the ELM as a broad theoretical perspective.

From an applied perspective, these data hold promise for their practical relevance in a variety of domains, including radio advertisements, judicial settings, healthcare, and of course, public speaking. For example, when selecting spokespersons, beyond evaluating their suitability based on features such as attractiveness, experience, and so forth, perhaps more attention should be given to how the speaker effectively modulates different properties of their voice when
communicating with others. Importantly, these data help identify several parameters of voice that play an important role in successfully persuading others, why they are effective, and the different thought processes that listeners may use when evaluating the quality of a spokesperson’s message. Thus, one can imagine how this information might be useful with respect to training effective salespersons as well as coaching individuals who may suffer from a fear of public speaking. Taken together, from both a practical and theoretical standpoint, these data are encouraging for their utility in advancing our understanding of the role played by vocal qualities in persuasive communication.

**Future Directions**

In considering future research, one possibility might be to investigate the processes by which vocal confidence affects persuasion at other levels of the elaboration continuum. Specifically, in our studies, we focussed on high and low levels of elaboration. However, the ELM postulates a different process when people are in the middle of the elaboration continuum. Here the ELM suggests that the primary effect vocal confidence should have under moderate elaboration is to influence the amount of processing. One might expect in this context that confidence should influence a listener’s motivation to attend to the message because confidence may reflect the extent to which a speaker is providing valuable/accurate information. Importantly, we would not necessarily expect all vocal properties to function in the same way. For example, whereas rate of speech and intonation should both affect perceptions of confidence and thus influence motivation to process a message in a similar manner, ability to process may be influenced in different ways. Specifically, extremely fast or slow rates of speech may negatively impact the listener’s ability to process a message. However, there is no compelling basis to suggest that variability in vocal intonation should necessarily affect ability. Testing these
predictions would provide a more comprehensive picture of the roles played by different properties of voice in persuasive communication.
References


Footnotes

1. The ELM proposes three ways a variable can influence attitudes under high-elaboration. Although we focused on the role of confidence as a biasing factor, the ELM suggests variables can also serve as an argument for or against an advocacy or in a self-validating role. Speaker confidence could serve as an argument if communicating with confidence was directly relevant to the advocacy, such as advertising a course on public speaking. Variables that serve in a self-validating role by enhancing or reducing confidence in one’s thoughts often have the greatest impact on thoughts when they are introduced after the message has been received. Thus, because how a speaker communicates is inherently part of the message, speaker confidence is unlikely to serve in a self-validating role.

2. As sample size relates to issues of power, the analyses most sensitive to sample size were the multi-sample SEM conducted in studies 1 and 2. A minimum of 10 participants per free parameter is recommended. Given that our multi-sample SEM analyses for study 1 had 15 free parameters in both high- and low-elaboration groups, we required at least 150 participants per group. Our multi-sample SEM for study 2 had 12 free parameters in both groups, thus requiring at least 120 participants per group. We continued data collection for the remainder of the academic term once those values were exceeded.

3. Participants indicated the fast-talking speaker ($M = 4.74$, $SE = .08$, 95% CI 4.59, 4.90) spoke faster than the slow-talking speaker ($M = 3.53$, $SE = .08$, 95% CI 3.37, 3.69), $F(1, 386) = 116.51$, $p < .001$, partial $\eta^2 = .23$. No main effect of vocal intonation emerged, $F(1, 386) = 3.29$, $p = .07$, partial $\eta^2 = .01$. However, a main effect of elaboration revealed that participants assigned to high-elaboration ($M = 4.30$, $SE = .08$, 95% CI 4.14, 4.45), rated the speaker as talking faster than low-elaboration ($M = 3.97$, $SE = .08$, 3.82, 4.13), $F(1, 386) = 8.37$, $p < .01$, partial $\eta^2 = .02$. A significant intonation by elaboration interaction, $F(1, 386) = 4.02$, $p = .046$, partial $\eta^2 = .01$, revealed that high-elaboration participants indicated the speaker spoke faster when the speaker’s intonation rose ($M = 4.51$, $SE = .11$, 95% CI 4.29, 4.74), versus fell ($M = 4.08$, $SE = .11$, 95% CI 3.87, 4.30), $p < .01$. Under low-elaboration, no difference in ratings of speech rate emerged across levels of intonation, $p = .89$. No further effects reached significance. Importantly, the absence of a speech rate by elaboration interaction suggests that constraining processing did not impact participant’s ability to detect changes in the speaker’s rate of speech. Thus, these data suggest that our rate of speech manipulation was successful.

4. In study 1 and 2, participants also rated the favorability of their thoughts. Analyses based on favorability ratings provided by participants produced similar but slightly weaker results as analyses based on favorability ratings provided by independent coders.

5. In study 1 and 2, using the total number of topic-relevant thoughts as the dependent variable produced a highly similar pattern.

6. In study 1 and 2, these relationships were also evaluated using a regression-based moderated mediation analysis. The results closely resembled those produced through multi-sample structural equation modelling.

7. Because all variables are measured variables and do not have multiple indicators, this model was specified such that each variable was fixed with a factor loading of 1 on its underlying construct and the associated error for each variable was set to 0. This assumes the measure is a perfect representation of its underlying construct. This assumption is implicitly made in all ANOVA and regression-based analyses and thus is not unique to our statistical approach. Additionally, our model presumed the residual variance in our endogenous variables were independent.

8. Participants rated the speaker with raised pitch ($M = 3.93$, $SE = .10$, 95% CI 3.74, 4.12) as having a higher pitched voice than the speaker with lowered pitch ($M = 2.54$, $SE = .10$, 95% CI 2.34, 2.73), $F(1, 408) = 102.25$, $p < .001$, partial $\eta^2 = .20$. No main effect of elaboration, $F(1, 408) = .85$, $p = .36$, partial $\eta^2 = .00$, or interaction between pitch and elaboration, $F(1, 408) = 2.46$, $p = .12$, partial $\eta^2 = .01$, was predicted and none emerged. The absence of a pitch by elaboration interaction suggests that constraining processing did not impact participant’s ability to detect changes in the speaker’s pitch. Thus, these data suggest that our vocal pitch manipulation was successful.
Figure 1.

*Path Model Depicting Relationships between Variables under High – and Low-Elaboration.*

Figure 2.

*Vocal confidence and cognitive responses as mediators of the relationship between vocal speed and vocal intonation and attitude.*

Panel A) High-Elaboration:

Panel B) Low-Elaboration:
Figure 3.

Vocal confidence and cognitive responses as mediators of the relationship between vocal pitch and attitude.

Panel A) High-Elaboration:

Panel B) Low-Elaboration: