**How Speaking Rate Shapes Consumer Response**

**Abstract**

From salespeople and customer service representatives to doctors and politicians, marketplace actors often communicate with their voice. But might articulation rate (i.e., how quickly one speaks) shape the impact of their communication? And if so, how? While prior psychological research suggests that speaking more slowly can sometimes be detrimental, in the context of social interactions, we suggest that the opposite may be true. Consistent with this suggestion, a multimethod investigation, including automated audio analysis of hundreds of real customer service calls and controlled experiments, demonstrates that speaking more slowly (within a range of normal speaking speed) boosts customer satisfaction and leads communicators to be perceived more positively. These effects are driven by perceived empathy. Speaking more slowly makes communicators seem more empathetic, which has positive downstream effects. Taken together, these findings shed light on articulation rate’s impact, deepen understanding around drivers of empathy, and highlight how automated audio analysis can provide insight into consumer behavior.

*Keywords*: articulation rate, empathy, automated audio analysis, social interactions, vocal features, person perception

Effective communication is an integral part of almost every marketplace interaction. It shapes how salespeople sell products, customer service representatives solve problems, doctors explain diagnoses, and politicians pitch policies. Indeed, recent work finds that subtle shifts in communicator language can impact everything from customer satisfaction and sales to word-of-mouth and engagement (e.g., Berger et al. 2020; Moore 2015; Moore and Lafreniere 2020; Packard and Berger 2024; Pogacar, Shrum, and Lowrey 2018).

But while it is clear that *what* communicators say (i.e., the words they use) shapes their effectiveness, might *how* they say it also play a role? Specifically, might the vocal features they use when communicating influence their impact?

To shed light on these questions, we focus on a vocal feature inherent in every spoken communication – articulation rate, or the rate at which people speak (Quené 2007). Communicators can speak more quickly or deliver words at a slower pace. While prior psychological research (e.g., Cesario and Higgins 2008; Ray 1986) suggests that speaking more slowly can sometimes be detrimental (because it signals a lack of confidence), in the context of social interactions, we suggest that the opposite may be true. Whether talking to customer service reps, doctors, or other marketplace actors, customers care about being heard. In these situations, we suggest that speaking more slowly (within a range of normal speaking speed) can boost customer satisfaction, and lead communicators to be perceived more positively, because it suggests that they are more empathetic (i.e., understand and care about the audience’s needs).

A multimethod investigation, combining automated audio analysis of hundreds of real customer service calls with controlled experiments, tests these possibilities. The studies demonstrate that speaking more slowly can boost customer satisfaction (and communicator evaluations) and illustrate the underlying role of perceived empathy in driving these effects.

This work makes several contributions. First, we shed light on how vocal cues shape consumer behavior. While a burgeoning stream of research has begun to investigate the effect of different linguistic features (e.g., words, phrases, or styles of language; Luangrath, Peck, and Barger 2017; Luangrath et al. 2023; Moore 2012; Moore and Lafreniere 2020; Patrick and Hagtvedt 2012a; Sela, Wheeler, and Sarial-Abi 2012), there has been less attention to vocal cues (c.f., Wang et al. 2021). We fill this gap, demonstrating how articulation rate shapes communication’s impact, and the underlying process that drives this effect.

Second, more narrowly, we provide an important corrective to existing literature. While some research suggests that speaking slower can be detrimental, it focused on monologues, or non-social situations where people listened to a recording of someone reading a book or dictating an opinion (e.g., Apple, Streeter, and Krauss 1979; Cesario and Higgins 2008). In the context of social interactions (e.g., customer service and doctor-patient interactions), however, we suggest that speaking more slowly can actually be beneficial because it makes communicators seem more empathetic.

Third, we deepen understanding around the drivers of empathy. Empathy is fundamental to social interactions (Hall and Schwartz 2019), shaping things like satisfaction (Singh and Sirdeshmukh 2000), purchase (Packard, Moore, and McFerran 2018), and prosocial behavior (Bagozzi and Moore 1994). But what leads communicators to be seen as more empathetic in the first place? We demonstrate that speaking more slowly signals that communicators understand and care about their audience’s needs.

Fourth, from a methodological standpoint, we illustrate how automated audio analysis can provide insight into consumer behavior. Recent work has highlighted the value of audio processing for marketing research (Grewal, Gupta, and Hamilton 2021), and begun to apply this approach to speech (Wang et al. 2021). But while there has been interest in voice data, measurement has proved difficult. Vocal features (e.g., pitch, volume, and articulation rate) occur simultaneously, making it hard to isolate and measure them in an objective, automated, scalable way (Balducci and Marinova 2018; Hildebrand et al. 2020). To address this issue, we use cutting-edge audio processing algorithms, demonstrating how automated audio analysis can measure key features at scale, and opening new avenues for future research.

Finally, these findings have clear practical implications. From salespeople and politicians to educators and healthcare professionals, communicators want to increase their impact. Our results indicate that speaking more slowly is a simple way to achieve this goal.

Further, compared to more complex resource-intensive strategies (e.g., extensive training or incentives), adjusting articulation rate should be easier and potentially more cost-effective to implement.

**Language and Communication**

Language is fundamental part of effective communication. Marketplace actors engage in countless conversations each day, and use language to shape perceptions, build relationships, and influence others’ decisions (Berger et al. 2022; Moore 2012; Pogacar et al. 2022; see Packard and Berger 2024 for a review).

Consistent with its importance, a great deal of research has examined how the words communicators use can affect their impact. When service agents refer to themselves as “I” rather than “we,” for example, it improves customer satisfaction because it makes customers feel like the agent cares and is more involved in the interaction (Packard, Moore, and McFerran 2018). Similarly, swear words can make reviews more helpful and boost product evaluations (Lafreniere, Moore, and Fisher 2022) and refusing things by saying “I don’t” rather than “I can’t” can enhance persuasion (Patrick and Hagtvedt 2012b).

But beyond *what* communicators say, might the *way* they say it (i.e., the vocal features used) shape the impact of their communication?

**Empathy**

We suggest this possibility based on research on empathy. Empathy is commonly described as one’s ability to understand and share the concerns of another (Davis 1994). It can involve seeing things from others’ point of view, imagining oneself in their place, or even feeling what someone else is feeling.

Not surprisingly, empathy plays a key role in social interactions (Hall and Schwartz 2019). Whether calling customer service, talking to salespeople, or seeking help from doctors, people want to be listened to and understood. Consequently, the more empathetic communicators seem, the more favorably audiences tend to respond (Eisenberg and Miller 1987). Seeing employees or public representatives as empathetic makes consumers more satisfied with services (Singh and Sirdeshmukh 2000), for example, more likely to buy products (Packard, Moore, and McFerran 2018), and more likely to adopt prosocial behaviors (Bagozzi and Moore 1994).

Perceptions of empathy are influenced by different cues. Mirroring another’s facial expression, for example, suggests that someone is experiencing the same emotions as their interaction partner (Preston and De Waal 2002). Similarly, employees’ use of personal pronouns (e.g., *I*’ll rather than *we*’ll solve your problem) can indicate more personal concern about a situation, and thus increase perceived empathy (Packard, Moore, and McFerran 2018). And gestures depicting train of thought (e.g., pointing and shrugging), concrete language, and providing feedback are all cues that signal listening (Ames, Maissen, and Brockner 2012; Burgoon, Guerrero, and Manusov 2011; Packard and Berger 2021).

But beyond words, or body gestures, might features of communicator’s voice also shape perceived empathy, and thus consumer response?

**Vocal Cues**

We suggest that the articulation rate, or the speed at which one speaks excluding pauses (Quené 2007), can shape how empathetic communicators seem. Vocal features (e.g., articulation rate, volume, and pitch) are an integral part of spoken communication. Marketplace actors can speak faster or slower, for example, softer or louder, and using a higher or lower pitch.

Psychological research has begun to examine how vocal features shed light on communicator’s internal states (Schroeder and Epley 2015) or how others perceive them. High pitch, for example, is linked to excitement (Laukka et al. 2016) and increased volume suggests that communicators are happy or scared (Juslin and Laukka 2003). Similarly, politicians who speak with a high pitch are perceived as having weaker leadership skills (Klofstad et al. 2012) and using low-pitched voices can signal calm and competence (Guyer et al. 2019; Oleszkiewicz et al. 2017).

But while scholars in psychology, communications, and linguistics have started to explore how vocal cues might offer insights into the communicator’s state, or shape how others perceive them, almost no research in marketing has explored vocal features’ impact. Indeed, while everyone from salespeople and customer service representatives to doctors and politicians communicate with their voice, to the best of our knowledge, only one paper in marketing has examined vocal cues. Wang et. al. (2021) focused on vocal tone in Kickstarter campaigns, and found that vocal tones denoting focus, low stress, and stable emotions encouraged funding.

**Articulation Rate**

To begin to explore how other vocal cues might shape important marketing outcomes, we examine articulation rate. Existing psychological work suggests that speaking more slowly can be detrimental. When listening to recordings of strangers talking about college admissions quotas, for example, Apple et al. (1979) found that people perceived others who spoke more slowly as more hesitant. Similarly, when listening to recordings of strangers talking about the clinical consequences of caffeine, Miller et al. (1976) found that people perceived others who spoke more slowly as less trustworthy. More broadly, whether listening to recordings of strangers talking about a school policy (Guyer et al. 2019), reading a book passage (Ray 1986), or explaining a school assistance program (Cesario and Higgins 2008), people perceived slower speakers as less competent and less confident. Taken together, this work suggests that speaking more slowly should lead communicators to be perceived more negatively and hurt marketing outcomes like customer satisfaction with service interactions.

In contrast, we suggest the opposite. Specifically, in the context of social interactions (e.g., with a customer service representative or doctor), we suggest that speaking more slowly can actually be beneficial.

Importantly, work suggesting negative effects of speaking slowly focused on non-social monologues. Participants were asked to listen to recordings of someone reading something (e.g., a book passage) or expressing their opinion to a tape recorder, and rated the speaker based on that. In situations like these, speaking more slowly may be detrimental because of the inferences it encourages. Speaking too slowly when reading a passage, for example, may make it seem like the speaker is uncertain, isn’t very confident, or doesn’t know what they are talking about (e.g., Cesario and Higgins 2008). Not surprisingly, then, in these types of contexts, speaking slowly may have negative effects.

In the context of social interactions, however, we suggest that speaking more slowly may have beneficial effects because it makes communicators seem more empathetic. Unlike other vocal cues, prior work often links articulation rate to speaker’s thoughtfulness (Miller, Grosjean, and Lomanto 1984; Quené 2007). Said simply, speaking more slowly may indicate the communicator is taking more time to think (Smith and Clark 1993). But while taking more time to think may be interpreted as lack of confidence when passive speakers read a book passage, in social interactions, speaking more slowly may suggest something quite different. In particular, it may generate the attribution that communicators are choosing their words carefully, mulling over what their conversation partner said, making sure they understand it, and addressing any concerns raised.

Indeed, empathy lies in understanding another person’s thoughts and feelings (Decety and Jackson 2004), and thoughtfulness serves as a key signal (Davis 1994). Just as taking more time to make a decision is associated with being more thoughtful and considerate (Kupor et al. 2014), speaking slower should suggest that the communicator is more empathetic (i.e., that they care about the audience’s needs and are taking the time to make sure they understand and respond to them). This, in turn, should make interaction partners feel more satisfied and perceive the speaker more favorably.

Note that inferences about empathy should be particularly likely in social interactions (compared to monologues) because this is a context where empathy thrives (Baston et al. 2022; Grühn et al. 2008). Indeed, whether it’s social conflict (Eisenberg and Lennon 1983), family interactions (Davis 1994), or organizational social dynamics (Dutton et al. 2006), empathy is key to resolve problems and foster connectedness. When listening to a recording of a stranger reading a book passage, or sharing their opinion about coffee, audiences should be unlikely to spontaneously consider how empathetic the speaker is, as the speaker is not responding to them. When interacting with a customer service agent or doctor, however, audiences are more likely to be attuned to whether speakers seem to care about them, and, as a result, make inferences about empathy.

**The Current Research**

Taken together, we suggest that speaking more slowly (within a normal range) should increase customer satisfaction and boost evaluations of the communicator. Further, we suggest that this should be driven by perceived empathy. Speaking more slowly should suggest speakers understand and care about the audience’s needs, which, in turn, should have positive downstream effect.

A multimethod approach tests these predictions. To provide preliminary evidence of articulation rate’s potential impact, Study 1 looks to the field. We use automated audio mining to analyze hundreds of customer service calls, examining whether customers are more satisfied when agents speak more slowly. In addition, to test alternative explanations, we account for a variety of additional features (e.g., aspects of the agent, customer, call, and things discussed) and see whether they can explain the effects.

Then, to more directly test articulation rate’s causal impact, and the underlying process, the next three studies use controlled experiments. Study 2 manipulates articulation rate, and measures the hypothesized process, demonstrating that speaking more slowly increases satisfaction because it signals empathy. Study 3 explores generalizability to a different context (i.e., doctor-patient interaction), voice, speaking rates, and outcomes. Study 4 further tests the process through moderation. If articulation rate’s effects are driven by perceived empathy, as we suggest, then they should be attenuated when people’s need for empathy is lower to begin with.

**Study 1: Articulation Rate in the Field**

To provide preliminary evidence of articulation rate’s potential impact, Study 1 looks to the field. We used automated audio analysis to measure how quickly employees spoke in hundreds of customer service calls, testing whether customers are more satisfied when employees use a slower articulation rate (i.e., speak more slowly).

***Method***

We worked with a large American online retailer to acquire recordings of 200 customer service phone calls. Each involves a different customer interacting with one of 129 customer service employees. Eleven calls were at least partially inaudible, leaving 189 recordings. Conversations last 5.06 minutes on average (*SD* = 3.63 minutes).

A professional transcription service converted the audio recordings to text. A research assistant broke each recording into conversational turns (e.g., employee turn 1: “Unfortunately we do not do price matching,” customer turn 2: “Okay, I know you used to,” employee turn 3: “Yeah, we used to”). The average conversation contains 67.88 turns (*SD* = 47.13), resulting in 12,830 turns overall.

Automated audio analysis was used to measure employee articulation rate. Following prior work (e.g., De Jong and Wempe 2009), articulation rate was operationalized as syllables per second of speaking time (excluding pauses). To do so, following the latest advancements in speech analysis and audio processing (e.g., Harma Saeed 2023; Hema and Marquez 2023; Morgan and Braasch 2022), we used *librosa* Python package (McFee et al. 2015). First, we used the Short-Time Fourier Transform (which breaks down the audio signal into a series of short-time frames) to convert the audio signal into a spectrogram (i.e., a 2D array where each row represents a different frequency band, and each column represents a different point in time). Figure 1, for example, illustrates the spectrogram derived from the agent’s phrase, “Very well, what can I do for you today?” The color intensity presents the amplitude (or energy) of different frequencies at different times. The redder (or hotter) the color, the higher the amplitude of those frequencies at that time; the bluer (or cooler) the color, the lower the amplitude.

Figure 1: Example Spectrogram

Immagine che contiene testo, schermata, Policromia, arte

Descrizione generata automaticamente

Second, we applied log compression to the spectrogram (conversion in decibels) to ensure that our measurement is more closely aligned with how humans perceive sounds.[[1]](#footnote-1)

Third, we identified where new syllables are likely to start. Syllables typically exhibit spectral patterns (i.e., unique blends of frequencies and intensities that define a certain sound) due to changes in vocal tract resonances, formants, and articulatory movements (Nittrouer and Lowenstein 2008). Consequently, we computed the spectral flux (i.e., a measure of how quickly the frequency content of the sound changes over time) to capture rapid changes that may indicate the start of a new syllable (i.e., “onsets”). To illustrate, the black line in Figure 2 shows the frequency trend of the voice, and the dashed vertical bars align with local peaks, or instances of rapid frequency changes. Each corresponds to a syllable within the phrase, “Very well, what can I do for you today?”.

Fourth, a *librosa* function identified peaks in the spectral flux where different syllables began and counted the number of syllables. Finally, articulation rate (*M* = 5.22, *SD* = .59) was computed as the number of syllables divided by seconds of speech. In Figure 2, for example, the algorithm identified a total of 11 syllables, resulting in an articulation rate of 7.34 syllables per second. Ancillary analysis demonstrate that this measure is highly correlated (*r* = .79) with human perceptions of speaking speed, underscoring its validity.[[2]](#footnote-2)

Figure 2: Example Syllables Detection Immagine che contiene linea, Diagramma, diagramma, schermata

Descrizione generata automaticamente

To capture customer satisfaction, we used a measure provided by the retailer. Perceived helpfulness is a key aspect of customer satisfaction (Cronin and Taylor 1992; Parasuraman et al. 1991), and at the end of each call, customers rated how helpful they felt the employee was (1 = not at all helpful to 4 = very helpful). For each of the four levels, the retailer provided a random sample of 50 calls.

Finally, ordinary least squares regression examined the relationship between articulation rate and customer satisfaction (results are also robust to ordered logistic regression, see robustness checks below). Given that the different variables do not share similar scales, all continuous variables were standardized (*z*-scored). Articulation rate was reverse coded so that positive coefficients reflect a positive relationship between speaking slower and customer satisfaction.

***Results***

As predicted, customers were more satisfied when employees used a slower articulation rate (i.e., spoke more slowly, *b* = .161, *SE* = .081, *t* = 1.99, *p* = .048; Table 1, model 1).

*Control Variables* While this initial result is intriguing, one could wonder whether it is driven by other factors. Consequently, we control for a range of alternative explanations, including aspects of the call, employee, and customer.

*Aspects of the Call.* Rather than articulation rate, one could argue that the results are driven by call attributes. Maybe employees speak slower for issues where customer satisfaction is higher, and perhaps this drove the results. We address this possibility in three ways. First, we control for six categories of *call reasons* (classified by the company as: account, gift card, order, product, return, and shipping). Second, we control for the topics discussed in each call using topic modeling (latent Dirichlet allocation, Blei 2012). We identified the lowest number of topics that maximize predictive power (i.e., 13 topics as identified by perplexity), and control for the proportion of each topic in each call (*call topics*). Third, we control for *call severity* asking two research assistants to code the severity of the issue discussed (1 = not at all, 5 = severe; Intraclass Correlation Coefficient ICC (2,1) = .720).

Beyond why customers are calling (the content of the issue), solving the issue should increase satisfaction, so we also control for whether the problem was solved or not (*solved issue*, identified by the company). Finally, longer calls might make it more difficult for employees to remain attentive (De Ruyter and Wetzels 2000), which can affect their performance and thus customer satisfaction, so we control for *call duration* (in seconds).[[3]](#footnote-3)

*Aspects of the Employee*. Beyond the call itself, one could wonder whether the results are driven by something about the employee. Part of this could be about employee characteristics. Experienced employees or particular employee demographics, for example, may increase customer satisfaction, so we control for *employee* *tenure* (in days) and *employee female* gender.

Alternatively, beyond the general topics of the call (mentioned above) other aspects of employee language could explain the results, so we control for this in several ways. First, concrete language can increase customer satisfaction (Packard and Berger 2021), so we control for *employee concreteness* using concreteness scores from Paetzold and Specia (2016). Second, questions can enhance liking (Huang et al. 2017), so we control for the proportion of *employee questions* using the Linguistic Inquiry and Word Count (LIWC; Pennebaker et al. 2015). Third, we account for other major language features controlling for the LIWC psychological process dictionaries (*employee* *psychological processes*: affect, bio, cognition, drives, informal, perception, relative, social, time orientation).

How employees speak could also explain the results, so we account for several other vocal features. First, more frequent pausing can increase customer satisfaction (Van Zant et al. 2023), so we control for *employee pauses* by counting the number of times the employee paused for at least 0.3 seconds.[[4]](#footnote-4) Second, speakers with low-pitched voices are perceived as more empathic (Apple, Streeter, and Krauss 1979), so we control for *employee pitch* using the YIN frequency estimator algorithm (De Cheveigné and Kawahara 2002)*.*[[5]](#footnote-5)

*Aspects of the Customer*. Beyond call or employee attributes, we also account for aspects of the customer. Part of this could be about customer characteristics, so we control for this in several ways. First, we control for the length of time (in days) customers had an account with the retailer prior to the call (*customer tenure*). Second, we control for customer gender (*female customer*), age (*customer age*, classified by the company as: 18-25, 25-34, 35-44, 45-54, 55+), and geographic region (*customer region*, classified by the company as: East, West, Midwest, South, Other).

We also control for various aspects of what customers say and how they say it. Customer articulation rate may lead employees to speak similarly in response (e.g., Buller, et al. 1992), which may affect customer satisfaction, so we control for that (*customer articulation rate*). In addition, we control for all the customer language (*customer concreteness*, *customer questions*, *customer psychological processes*) and paralanguage (i.e., *customer pauses* and *customer pitch*) features we included for employees.

*Results Including Controls*. Even after accounting for all these controls, customers were still more satisfied when employees spoke slower (*b* = .182, *SE* = .080, *t* = 2.27, *p* = .025; Table 1, model 2). [[6]](#footnote-6)

Table 1: Articulation Rate and Satisfaction

|  |  |  |
| --- | --- | --- |
|  | (1) Base Model | (2) With Controls |
| **Slower Artic. Rate**  Controls  *Call*  Call Reasons  Call Topics  Call Severity  Solved Issue  Call Duration  *Employee* | **.161\*\* (.081)** | **.182\*\* (.080)**  Included  Included  Included  Included  Included  Included |
| Tenure |  | Included |
| Female |  | Included |
| Pauses  Pitch  Concreteness  Questions  Psych. Processes |  | Included  Included  Included  Included  Included |
| *Customer*  Tenure  Female  Age  Region  Artic. Rate  Pauses  Pitch  Concreteness  Questions  Psych. Processes  R-squared | .021 | Included  Included  Included  Included  Included  Included  Included  Included  Included  Included  .600 |
| N | 189 | 189 |

*Notes*: \*\**p* < .05. Standard errors are in parentheses. Results for controls are in Table WA1.

***Robustness***

We also ran several robustness tests. First, one could wonder whether the results are driven by the particular articulation rate measure used. To test this possibility, we used an alternative approach. Specifically, following prior work (Šošić and Graovac 2022), we used the *pyphen* Python package to determine the number of syllables the employee used in a given call, and divided this by the speaking time. Even using this alternate way of measuring articulation rate, however, results remain the same: speaking more slowly increased customer satisfaction (*b* = .191, *SE* = .077, *t* = 2.47; *p* = .015; Table WA1, column 2).

Second, one could wonder whether the results are somehow driven by the modeling approach used. An ordinal logistic regression treating customer satisfaction as a four-level discrete outcome, however, finds the same results (*OR* = 1.923, *SE* = .421, *t* = 3.15, *p* = .002; Table WA1, column 3). Results also persist using Lasso penalized linear regression (*b* = .135, *SE* = .067, *t* = 2.01, *p* = .046; Table WA1, column 4) that eliminate non-essential controls and account for collinearity (Tibshirani 1996).

Third, one could argue that rather than bring driven by speaking speed, the effects are driven by response speed. Beyond how quickly someone talks *while* they are speaking, they can also respond more or less quickly to what someone else said. If people who speak more slowly also tend to respond more slowly, and responding slowly is seen as a signal of thoughtfulness,[[7]](#footnote-7) then maybe response speed (rather than speaking speed) is driving the effects. To test this possibility, we measured how long it took (in seconds) each agent to respond after each customer turn (averaged across turns). Articulation rate and response speed are weakly correlated (*r* = .11), however, and when response speed is included as a control in the full model, the effect of speaking slower persists (*b* = .181, *SE* = .079, *t* = 2.29, *p* = .024). Consequently, response speed cannot explain the effects observed here (the experiments cast further doubt on this alternative by controlling for response speed across conditions).

***Discussion***

Study 1 provides preliminary evidence for how articulation rate might shape customer satisfaction. Analyzing hundreds of customer service calls demonstrates that, consistent with our theorizing, customers were more satisfied when employees spoke more slowly. These results are robust to a variety of controls and model specifications.

Note that we are not suggesting that service agents should speak *extremely* slowly. Rather, we are simply suggesting that, within a normal range of speaking speeds, speaking more slowly can be beneficial. Indeed, the customer service representatives studied spoke with similar average speed (*M* = 5.22 syllables per second, *SD* = .59) as English-speaking adults (*M* = 5.12, Jacewicz et al. 2009). Further, even the slowest speaker (*M* = 3.63 syllables per second) did not speak extremely slowly. In the time it would take the average speaker to say “I appreciate your patience. Let me check that for you right away,” the slowest speaker would have been able to convey much of the same information (i.e., “I appreciate your patience. Let me check that…”).

One might also wonder whether articulation rate has a quadratic effect. Maybe speaking faster is beneficial up to the average speed (Rodero 2020) but speaking too quickly could backfire. Adding a quadratic term to the model, though, finds it is not significant (*b* = –.003, *SE* = .057, *t* = –.06, *p* = .956), casting doubt on this possibility. That said, to further test the idea that speaking more slowly is more beneficial within a normal range of speeds, Studies 3 and 4 directly compare slower articulation rate with the average speed.

**Exploring the Underlying Process**

While results of Study 1 are consistent with our theorizing, one could wonder what is driving the effect. While Studies 2-4 directly test the hypothesized underlying process through experiments, we begin to explore the underlying process (and alternative explanations) through ancillary analyses of the data from Study 1.

***Moderating Role of Empathy Language***

If a slower articulation rate increases satisfaction because it makes customers believe that the employee is more empathetic, as we suggest, then the effect should be mitigated when it’s already clear that the employee is empathetic. To test this possibility, we explored the moderating role of employee empathic language. Using Herhausen et al.’s (2023) empathy lexicon, developed in the context of customer service, we measured how many empathy words each employee used and test whether it moderates the effect.

A positive main effect of empathy language (*b* = .250, *SE* = .101, *t* = 2.47, *p* = .015) was qualified by the predicted articulation rate × empathy interaction (*b* = –.198, *SE* = .097, *t* = –2.04, *p* = .043, see Figure 3). Spotlight analyses one SD above and below the mean provide deeper insight into the pattern of results. Consistent with our suggestion about the underlying role of perceived empathy, when employees used less empathetic language (–1SD) to begin with, customers were more satisfied when employees spoke more slowly (*b* = .339, *SE* = .124, *t* = 2.72, *p* = .007). When employees already used more empathetic language (+1SD), though, consistent with our theorizing, articulation rate’s impact was mitigated (*b* = .057, *SE* = .125, *t* = .46, *p* = .648). In this case, employees should already be perceived as more empathetic, and thus articulation rate should have less of an impact.

Figure 3: Moderation by Empathy Language

***Testing Alternative Process Explanations***

While this moderation is consistent with our theorizing, one could still wonder about alternative explanations. Maybe speaking too quickly made it harder for customers to understand what agents were saying, and this, rather than perceived empathy, is driving the effects. To test this possibility, a research assistant rated each call on how easy it was to understand what the service representative said (1 = extremely easy to understand, 5 = extremely difficult to understand). Casting doubt on this alternative, however, articulation rate and understandability were only weakly related (*r* = .13), and even controlling for understandability, customers were still more satisfied when service representatives spoke more slowly (*b* = .182; *SE* = .083, *t* = 2.21; *p* = .030).[[8]](#footnote-8) Overall, this casts doubt on the notion that understandability could be driving the effect.

Alternatively*,* maybe the effects were somehow driven by mimicry (Moore and McFerran 2017). If customers who tend to be more satisfied speak more slowly, and service agents mimic customers’ speaking speed, this, rather than empathy, could be driving the results. To address this possibility, we conducted a turn-by-turn analysis using a vector auto-regression model. We lagged employee and customer articulation rate in each turn on their conversation partner’s prior turn(s). Casting doubt on this alternative, however, employee articulation rate was not impacted by customer articulation rate in the prior turn (lag 1 *b* = –.005; *SE* = .013, *t* = –.43; *p* = .671). Lags greater than one were also not significant. Further, customer articulation rate was not impacted by employee articulation rate in the prior turn (lag 1 *b* = –.002; *SE* = .015, *t* = –.14; *p* = .885). Overall, this casts doubt on the notion that mimicry could be driving the effect.

**Study 2: Directly Testing Causality and Process**

Results of Study 1 are consistent with our suggestion that speaking more slowly boosts customer satisfaction, but one could still wonder whether the relationship is truly causal. Controlling for various factors casts doubt on alternative explanations, but an even stronger test would be to experimentally manipulate articulation rate and measure its impact. Study 2 does this.

In addition, Study 2 further tests the hypothesized underlying process. While ancillary analyses of Study 1 support the notion that perceived empathy is driving the effects, Study 2 directly measures perceived empathy, and tests whether, as predicted, speaking more slowly increases satisfaction because it makes communicators seem more empathetic.

Finally, the study also tests alternative explanations. While the ancillary analyses of Study 1 cast doubt on the notion that understandability or mimicry could be driving the effects, Study 2 further tests understandability as well as conversationality, duration, and confidence. In addition, to further test whether response speed[[9]](#footnote-9) can explain the effects, we control for it across experimental conditions.

***Method***

Participants (*N* = 456; Prolific) were randomly assigned to a condition in a 2 (articulation rate: slow vs. fast) between-subjects design. See web appendix for exclusions, demographics, materials, and further details for all experiments.

Participants were asked to imagine calling customer service for a billing issue. Then, they listened to the agent’s response through one of two professional voice recordings (see web appendix for script and audio recordings).[[10]](#footnote-10)

The only difference between conditions was how slowly the agent spoke. In the slow condition, the agent spoke one standard deviation slower than the average speed of agents in Study 1 (i.e., 4.63 syllables/second). In the faster condition, they spoke one standard deviation faster (i.e., 5.81 syllables/second). A professional voice actor recorded both conditions and went through a week of preparation before recording. Further, to ensure they spoke at the desired speed, a video, generated from a karaoke synchronized lyrics program, guided them through each line of the script at the specified pace of each condition. This also controlled for response speed across conditions. The speaker was instructed to maintain the same pitch, intonation, loudness, pausing, and emphasis across recordings, and indeed, audio analysis using the *librosa* Python package revealed no differences on these dimensions across conditions (i.e., pitch: 46.8 vs. 46, intonation: .18 vs. .17, loudness: .02 vs. .03, and pausing: 0 vs. 0). In addition, confirming the manipulation’s effectiveness, a pretest indicated that the slow condition was perceived as slower (*M*slow = 3.00 vs. *M*fast = 5.45, *F*(1, 58) = 85.38, *p* < .001, η2 = .595).[[11]](#footnote-11)

Next, we measured the dependent variable. Following Study 1, participants rated customer satisfaction (i.e., how helpful they thought the agent was; 1 = not at all helpful, 7 = very helpful).

Then, we measured the hypothesized process. Using a 3-item scale adapted from prior work (Packard, Moore, and McFerran 2018), participants rated the degree to which they thought the agent was empathic (i.e., “empathic,” “understanding” and “concerned,” α = .92).

Finally, participants completed a manipulation check, ancillary measures to test alternative explanations (i.e., understandability, conversationality, duration, and confidence; more details below), two attention checks, and demographics.

***Results***

*Dependent Variable*. As predicted, and consistent with Study 1, speaking slower increased customer satisfaction (*M*slow = 5.97 vs. *M*fast = 5.70, *F*(1, 454) = 4.85, *p* = .028 η2 = .010).

*Hypothesized Mechanism*. In addition, consistent with our theorizing, speaking slower made the agent seem more empathetic (*M*slow = 5.33 vs. *M*fast = 4.90, *F*(1, 454) = 12.27, *p* = .001, η2 = .026).

*Mediation*. Finally, as expected, mediation (PROCESS model 4; Hayes 2018) found that perceived empathy mediated the effect of articulation rate on satisfaction (indirect effect = .27, 95% CI = .11, .45). Speaking more slowly made the agent seem more empathetic, (*b* = .44, SE = .12, *t* = 3.5, *p* < .001), which increased satisfaction (*b* = .63, SE = .04, *t* = 17.55, *p* < .001). Further, including empathy as mediator led the direct effects to be reduced to non-significance (*b* = –.01, 95% CI = –.19, .19), indicating “full” mediation.

*Alternative explanations*. Ancillary measures cast doubt on several alternative explanations. First, rather than being driven by perceived empathy, maybe speaking slower signaled confidence. To test this possibility, participants rated how confident the agent seemed (1 = not at all confident, 7 = very much confident). Casting doubt on this alternative, however, there was no effect of condition (*F*(1, 454) = .92, *p* = .337, η2 = .002).

Second, maybe the effect was driven by expectations. Maybe the fast speed went faster than participants expected it should, which made them feel negatively. To test this possibility, we asked participants to rate how the duration of the conversation was compared to what they would have expected (1 = not at all long, 7 = very much long). Duration did not vary by condition (*F*(1, 454) = .05, *p* = .820, η2 < .001), however, casting doubt on this alternative.

Third, maybe speaking more slowly sounded more conversational. To test this possibility, we asked participants how conversational the recording sounded (1 = not at all conversational, 7 = very much conversational). Conversationality did not vary by condition (*F*(1, 454) = 2.53, *p* = .126, η2 = .005), however, casting doubt on this alternative.

Fourth, maybe the faster articulation rate was harder to understand, and that drove the effects. To test this possibility, we asked participants how difficult it was to understand what the agent was saying (1 = extremely easy to understand, 7 = extremely difficult to understand). Unlike Study 1, here the faster condition was seen as harder to understand (*F*(1, 454) = 24.79, *p* < .001, η2 = .052), but adding understandability to parallel mediation with empathy indicated that empathy continued to mediate (*b* = .27, 95% CI = .11, .44) whereas understandability did not (*b* = .04, 95% CI = –.01, .09). This casts doubt on the notion that understandability is driving the effects, and we further test this point in Study 3.

***Discussion***

Study 2 provides direct causal support for our theorizing while also illustrating a mechanism behind the effects. First, consistent with Study 1, articulation rate shaped customer response. In this case, speaking more slowly boosted customer satisfaction.

Second, consistent with our theorizing, these effects were driven by perceived empathy. Speaking more slowly made the communicator seem more empathic, which boosted satisfaction.

Third, the study casts doubt on a range of alternative explanations. While one could wonder whether understandability, conversationality, duration, and confidence drove the effects, none of these factors can explain the result. Further, by controlling for response speed across conditions, Study 2 casts further doubt on this alternative.

**Study 3: Exploring Generalizability**

While the results of Study 2 are consistent with our theorizing, one might wonder whether they are somehow restricted to the domain, stimuli, or dependent variables examined. Consequently, Study 3 explores generalizability in several ways. First, to explore whether the effects generalize beyond customer service, we explore a different domain (i.e., doctor-patient interactions). Second, we use a different speaker and different stimuli. Third, to examine whether the results extend to other dependent variables, measuring satisfaction with the interaction more generally, as well as the efficacy of the communicator. Fourth, we again measure understandability and the other alternatives used in Study 2 to see whether they can explain the effects.

In addition, we also test another alternative explanation. While Study 2 found that speaking slower was more effective, one could argue that the effects are driven by the downsides of speaking too quickly, rather than the benefits of speaking slowly. Consequently, to more clearly test the benefits of speaking more slowly, we compare speaking slowly to speaking at an average speed.

***Method***

Participants (*N* = 362; Prolific) were randomly assigned to a condition in a 2 (articulation rate: slow vs. average [control]) between-subjects design.

Participants were asked to imagine visiting a doctor for a persistent and distressing medical condition. Then, they listened to the doctors’ response through one of two professional voice recordings. The recordings were made by a (different) professional voice actor and followed the training protocols of Study 2 to ensure strong experimental control.[[12]](#footnote-12)

The only difference between conditions was how slowly the “doctor” spoke. In the control condition, they spoke at the average speed of agents in Study 1 (i.e., 5.22 syllables/second), while in the slower condition, they spoke one standard deviation slower (i.e., 4.63 syllables/second). Confirming the manipulation’s effectiveness, a pretest indicated that the slower condition was perceived as slower (*M*slow = 3.88 vs. *M*average = 4.52, *F*(1, 118) = 14.73, *p* < .001, η2 = .111) but equally easy to understand (*M*slow = 1.75 vs. *M*average = 1.95, *F*(1, 118) = .58, *p* = .449, η2 = .005).

Next, participants completed the dependent variables. In addition to the satisfaction measure from Study 2, participants provided their satisfaction with the interaction (“how satisfied are you with the doctor’s response?” and “how satisfied are you with your experience with the doctor so far?”; 1 = not at all satisfied, 7 = very much satisfied, *r* = .92; Packard and Berger 2021), and how good they thought doctor was (1 = very bad, 7 = very good).

Then, they rated perceived empathy using the same measure as Study 2 (α = .92).

Finally, participants completed a manipulation check, the alternative explanation measures from Study 2, two attention checks, and demographics.

***Results***

*Dependent Variables*. As predicted, and consistent with the prior studies, speaking slower increased customer satisfaction (*M*slow = 6.03 vs. *M*average = 5.67, *F*(1, 359) = 9.99, *p* = .002, η2 = .027). It also increased participants’ satisfaction with the interaction overall (*M*slow = 5.96 vs. *M*average = 5.67, *F*(1, 359) = 5.66, *p* = .018, η2 = .015) and made the doctor seem like a better doctor (*M*slow = 6.06 vs. *M*average = 5.76, *F*(1, 359) = 6.31, *p* = .012, η2 = .017).

*Perceived* *Empathy*. In addition, consistent with our theorizing and Study 2, speaking slower made the doctor seem more empathetic (*M*slow = 5.94 vs. *M*average = 5.56, *F*(1, 359) = 10.96, *p* = .001, η2 = .030).

*Mediation*. Finally, as expected, mediation (PROCESS model 4; Hayes 2018) found that perceived empathy mediated the effects of speaking speed. Speaking more slowly made the doctor seem more empathetic, (*b* = .38, SE = .11, *t* = 3.31, *p* = .001), which increased customer satisfaction (*b* = .79, SE = .03, *t* = 23.49, *p* < .001), satisfaction with the interaction (*b* = .87, SE = .03, *t* = 26.88, *p* < .001), and boosted evaluations (*b* = .84, SE = .03, *t* = 25.13, *p* < .001). The resulting 95% confidence interval indicated significant indirect effects for customer satisfaction (*b* = .30, 95% Confidence Interval (CI) = .12, .49), satisfaction with the interaction (*b* = .33, 95% CI = .13, .52), and evaluation (*b* = .31, 95% CI = .12, .52), and including empathy as mediator led the direct effects to be reduced to non-significance (customer satisfaction: *b* = .07, 95% CI = –.08, .21; satisfaction with the interaction: *b* = –.04, 95% CI = –.18, .10; evaluation: *b* = –.02, 95% CI = –.16, .13), indicating “full” mediation.

*Alternative Explanations*. Ancillary analyses also tested whether alternative explanations (i.e., understandability, conversationality, duration, and confidence)could explain the effects. There was no effect of condition on understandability (F(1, 359) = 1.76, p = .185, η2 = .005), conversationality (F(1, 359) = .33, p = .567, η2 < .001), confidence (F(1, 359) = .29, p = .590, η2 < .001), or duration (F(1, 359) = .99, p = .320, η2 = .003), however, casting further doubt on these alternatives.

***Discussion***

Study 3 provides further evidence for both articulation rate’s impact and the process underlying these effects. First, as predicted, and consistent with the first two studies, speaking more slowly increased customer satisfaction (i.e., made communicators seem more helpful). It also increased satisfaction with the interaction more broadly, and made the communicator seem like a better doctor.

Second, consistent with our theorizing, and the results of the first two studies, these effects are driven by perceived empathy. Speaking more slowly made communicators seem more empathic, which had positive downstream effects.

Third, the fact that these results hold using a different voice, context (i.e., doctor-patient interactions), and speaking speeds, speaks to their generalizability.

**Study 4: Process by Moderation**

Study 4 further tests the process through both mediation and moderation. If the benefits of speaking slower occur because it makes communicators seem more empathic, as we suggest, then its effect should be moderated by how much listeners are looking for empathy in the first place.

To test this possibility, in addition to manipulating articulation rate, we also manipulate participants’ need for empathy. If our theorizing about the role of empathy is correct, the effect of articulation rate should be mitigated when need for empathy is low to begin with.

***Method***

Following pre-registration (<https://aspredicted.org/W59_QNH>), the final sample consisted of 802 people randomly assigned to a condition in a 2 (articulation rate: slow vs. average) × 2 (need for empathy: baseline vs. low) between-subjects design.

The baseline condition was the same as in Study 3. In the low need for empathy condition, however, we adjusted the scenario to reduce the need for empathy. In particular, we asked participants to imagine they had been experiencing only a mild medical condition that the doctor informed them was easy to solve. Confirming the manipulation’s effectiveness, a pretest indicated the people in the low need for empathy condition were less likely to be looking for empathy (*M* = 4.52 vs. 5.32, *F*(1, 78) = 5.65, *p* = .020, η2 *=* . 067).[[13]](#footnote-13)

Dependent variables and process were the same as in Study 3. Participants then completed two manipulation checks, the attention checks used in Study 3, and demographics.

***Results***

*Customer* *Satisfaction*. Main effects of articulation rate (*F*(1, 798) = 3.73, *p* = .054, ηp2 = .005), and need for empathy (*F*(1, 798) = 9.70, *p* = .002, ηp2 = .012), were qualified by the predicted articulation rate × need for empathy interaction (*F*(1, 798) = 3.44, *p* = .064, ηp2 = .004; see Figure 4). Consistent with the prior studies, in the baseline condition, speaking slower increased customer satisfaction (*M*slow = 5.99 vs. *M*average= 5.67, *F*(1, 798) = 7.12, *p* = .009, ηp2 = .020). Consistent with the hypothesized underlying role of empathy, however, when people needed less empathy, the beneficial effect of speaking slowly was mitigated (*M*slow = 5.58 vs. *M*average= 5.57, *F*(1, 798) = .01, *p* = .957, ηp2 < .001). Further, this occurred because lower empathy needs reduced the positive effect of slow articulation rate on satisfaction (*M*baseline = 5.99 vs. *M*low need = 5.58, *F*(1, 798) = 12.38, *p* < .001, ηp2 = .034). Similar effects were observed for both satisfaction with the interaction and evaluations, so given space constraints, we report these results in the web appendix.

Figure 4: Need for Empathy Moderates the Effect

Error bars: +/– 1 SE

*Empathy*. Similar effects were found on perceived empathy. Main effects of both articulation rate (*F*(1, 798) = 7.90, *p* = .005, ηp2 = .010) and need for empathy (*F*(1, 798) = 13.00, *p* < .001, ηp2 = .016) were qualified by the predicted articulation rate × need for empathy interaction (*F*(1, 798) = 8.18, *p* = .004, ηp2 = .010). Consistent with Study 3, in the baseline condition, speaking slower made the doctor seem more empathetic (*M*slow = 5.85 vs. *M*average= 5.42, *F*(1, 798) = 16.00, *p* < .001, ηp2 = .046). Consistent with the hypothesized underlying role of empathy, however, when people’s need for empathy was lower, the beneficial effect of speaking slowly was mitigated (*M*slow = 5.36 vs. *M*average= 5.36, *F*(1, 798) = .01, *p* = .972, ηp2 < .001).

*Moderated Mediation*. Moderated mediation analysis (PROCESS model 8; Hayes 2018) found significant moderated mediation (satisfaction *b* = .32, 95% CI = .10, .54; satisfaction with the interaction *b* = .34, 95% CI = .10, .59; and evaluation *b* = .31, 95% CI = .09, .52). As in Study 3, in the baseline condition, the effect of speaking slower was driven by empathy (satisfaction: *b* = .31, 95% CI = .17, .46, satisfaction with the interaction: *b* = .34, 95% CI = .18, .50, evaluation: *b* = .31, 95% CI = .17, .46). Speaking slower made it seem like the doctor was more empathetic (*b* = .43, SE = .11, *t* = 4.00, *p* < .001), which increased customer satisfaction (*b* = .73, SE = .03, *t* = 25.35, *p* < .001), satisfaction with the interaction (*b* = .79, SE = .03, *t* = 27.85, *p* < .001), and evaluation (*b* = .71, SE = .03, *t* = 27.75, *p* < .001). When people’s need for empathy was lower, however, speaking slower no longer made the doctor see more empathetic (*b* = –.01, SE = .11, *t* = –.03, *p* = .972), and the mediation was no longer significant (satisfaction: *b* = –.01, 95% CI = –.17, .17, satisfaction with the interaction: *b* = –.01, 95% CI = –.19, .18, evaluation: *b* = –.01, 95% CI = –.16, .16).

***Discussion***

Study 4 underscores both the beneficial effects of speaking more slowly, and the underlying mechanism behind these effects. First, consistent with our prior studies, speaking more slowly boosted customer satisfaction and evaluations of communicators. In this instance, when the doctor spoke more slowly, they were seen as a better doctor, and patients were more satisfied.

Second, the results underscore the underlying role of empathy through both mediation and moderation. Speaking more slowly made the communicators seem more empathetic, which had beneficial downstream effects. Further, consistent with the notion that the effects are driven by perceived empathy, they were attenuated when people’s need for empathy was lower.

It’s worth noting that these results are consistent with the Study 1 ancillary analyses, even though the two paradigms examine empathy’s moderating role in slightly different ways. Study 1 examined *other communicator empathy* *cues*, and found that when other cues (i.e., communicator language) already signaled empathy, articulation rate’s effects were reduced. Study 4, in contrast, looks at *recipients’* *need for empathy*, and finds that articulation rate has a weaker effect when empathy is less needed. Taken together, the studies underscore the hypothesized underlying role of empathy.

**General Discussion**

From customer service and sales to healthcare professionals and politicians, many marketplace interactions involve the spoken word. But while it’s clear that *what* communicators say (i.e., the words, phrases, and linguistic features they use) impacts the effectiveness of their communication, less is known about how communicators’ vocal features might also play a role.

The present research begins to address this gap. A multimethod investigation, mixing field data with controlled experiments, demonstrates the impact of articulation rate (i.e., how quickly communicators speak) and the psychological mechanism underlying these effects. First, automated audio analysis of hundreds of customer service calls finds that customers are more satisfied when agents speak more slowly. Experimentally manipulating articulation rate (Studies 2-4) underscores articulation rate’s causal impact, illustrating that it boosts both customer satisfaction and communicator evaluations.

Second, results shed light on the underlying process through both mediation (Studies 2-4) and moderation (Studies 1 and 4). Speaking more slowly has beneficial effects because it makes communicators seem more empathetic. Furthermore, consistent with the notion that the effects are driven by perceived empathy, they are mitigated when there are other cues to empathy (i.e., speaker language, Study 1) or listeners have lower needs for empathy in the first place (Study 4).

Third, the studies cast doubt on various alternative explanations. The effects persist in the field even accounting for aspects of call, agent, customer, and interaction. Ancillary analyses also cast doubt on the possibility that understandability, conversationality, duration, confidence, or closeness are driving the observed effects.

Finally, the fact that the results hold across speaking speeds (i.e., fast vs. slow and average vs. slow), voices (i.e., male and female), outcome variables (i.e., customer satisfaction, satisfaction with the interaction, and communicator evaluations), and domains (i.e., employee-customer and doctor-patient interaction) speaks to their generalizability.

***Contributions and Implications***

This research makes several contributions. First, we highlight the importance of vocal features in shaping consumer behavior. A burgeoning stream of research has begun to examine how the specific words communicators use affect their impact (e.g., Moore and Lafreniere 2020; Patrick and Hagtvedt 2012a; Pogacar et al. 2022; Sela, Wheeler, and Sarial-Abi 2012). But while these studies have provided many valuable insights, less attention has been paid to the role of vocal cues. Our findings demonstrate that the articulation rate communicators use when speaking can have an important effect. Given that a broad range of interactions involve spoken communication, and the fact that the *way* people speak can be sometimes more impactful than the words they use (Van Zant and Berger 2020), this area is ripe for further attention.

Second, more narrowly, we provide an important corrective to the notion that speaking more slowly is detrimental. While some have suggested that speaking slower can lead communicators to be perceived more negatively, they focused on non-social settings and recordings of passive speakers (e.g., Cesario and Higgins 2008; Ray 1986). In the context of social interactions, however, we demonstrate that speaking more slowly can have the opposite effect. It increases customer satisfaction and boosts evaluations because it makes communicators seem more empathetic.

Third, we contribute to understanding what drives perceived empathy in the first place. While it is clear that being perceived as empathetic is beneficial for a range of social situations (Hall and Schwartz 2019), prior work has primarily investigated the impact of words or body language (e.g., Burgoon, Guerrero, and Manusov 2011; Packard and Berger 2021). Our research demonstrates that vocal cues can also have an important effect.

Fourth, methodologically, we highlight the value of using automated audio analysis in marketing research. While recent work has begun to adopt natural language processing and computer vision (e.g., Berger et al. 2020; Hartmann et al. 2021; He, Li, and Wang 2023; Humphreys and Wang 2018; Li and Xie 2020; Li, Shi, and Wang 2019; Marinova and Balducci 2018; Wang et al. 2022), there’s been less attention to auditory features (c.f., Wang et al. 2021). Given the prevalence of auditory attributes across contexts (e.g., interactions with people, firms, and technology) and forms of contemporary consumption (e.g., podcasts, TikTok videos, and songs), this area deserves deeper exploration.

Finally, the findings have important practical implication for marketplace actors. Frontline employees seek to build loyalty, politicians want to be trusted, and educators strive to foster growth. Our results suggest that a relatively simple shift, speaking slower, can help. In the context of customer service, for example, some call centers aggressively manage response time and cost per interaction. Our findings suggest they may want to potentially reconsider some of the potential consequences of these strategies.

***Directions for Future Research***

This work raises interesting questions for future research. First, individual differences might also play a role. People vary in their need for empathy (Davis 1994), and this might moderate articulation rate’s effects. Similarly, insecure individuals (Mikulincer et al. 2005), for example, or those who have had a bad prior experience in the domain (Davis 1983), might have greater needs for empathy, and thus articulation rate might have stronger effects.

Second, work might explore whether articulation rate’s effects are moderated by other verbal and nonverbal factors. Foreign accents can be seen as less familiar and empathetic (Wang et al. 2013), for example, potentially amplifying the articulation rate’ impact. Similarly, eye-contact and concrete language suggests active listening (Burgoon, Guerrero, and Manusov 2011; Packard and Berger 2021), and low-pitched voices are seen as more tender (Juslin and Laukka 2003), so articulation rate’s effects could be mitigated by these factors.

Third, automated audio analysis could be applied to a range of other domains. Consumers increasingly interact with AI voice-based systems (e.g., Alexa and Siri), for example, to get help, collect information, and buy products. Might the vocal cues they use predict their intentions? And if so, might using consumer vocal cues as input help AI systems provide better responses? People’s decisions often reflect their emotional state (Lerner et al. 2015), and how they speak is indicative of how they feel (Burgoon, Guerrero, and Manusov 2011). Higher pitch, for example, can suggest excitement (e.g., about a product; Juslin and Laukka 2003). Consequently, when systems detect consumers using higher pitch, they might want to offer a limited-time discount to encourage purchase. Similarly, changes in speech pattern (e.g., a pause) can signal that consumers are deliberating or evaluating information, so systems could provide additional details or recommendations to facilitate decision-making.

Beyond AI interaction, similar tools could be used to improve interactions with firms. Consumers frequently speak to firms (e.g., in customer service calls), and changes in vocal tone or cadence might indicate rising frustration (Burgoon and Bacue 2003; Juslin and Laukka 2003; Rochman, Diamon, and Amir 2008). Consequently, firms might use such shifts as signals that a different response is needed. Firms could also look for vocal markers of customer satisfaction and use that to determine when to initiate cross selling or upselling. Overall, audio mining can shed light on consumer-firm dynamics and uncover patterns that can help businesses refine marketing strategies.

***Conclusion***

In conclusion, the present research demonstrates that a subtle shift in how communicators speak can have important consequences for consumer perceptions and behavior. In so doing, it sheds light on the effects of vocal features in the marketplace, and on consumer behavior more broadly.

**References**

Ames, Daniel, Lily Benjamin Maissen, and Joel Brockner (2012), “The role of listening in interpersonal influence,” *Journal of Research in Personality*, 46 (3), 345-349.

Apple, William, Lynn A. Streeter, and Robert M. Krauss (1979), “Effects of pitch and speech rate on personal attributions,” *Journal of personality and social psychology*, 37 (5), 715.

Bagozzi, Richard P. and David J. Moore (1994), “Public service advertisements: Emotions and empathy guide prosocial behavior, “ *Journal of marketing*, 58 (1), 56-70.

Balducci, Bitty and Detelina Marinova (2018), “Unstructured data in marketing,” *Journal of the Academy of Marketing Science*, 46, 557-590.

Batson, C. Danile, Johee Chang, Ryan Orr, and Jennifer Rowland (2002), “Empathy, attitudes, and action: Can feeling for a member of a stigmatized group motivate one to help the group?,” *Personality and Social Psychology Bulletin*, 28 (12), 1656-1666.

Berger, Jonah, Ashlee Humphreys, Stephan Ludwig, Wendy W. Moe., Oded Netzer, and David A. Schweidel (2020), “Uniting the Tribes: Using Text for Marketing Insight,” *Journal of Marketing*, 84 (1), 1-25.

Berger, Jonah, Grant Packard, Reihane Boghrati, Ming Hsu, Ashlee Humphreys, Andrea Luangrath, Sarah Moore, Gideon Nave, Cristopher Olivola, and Matthew Rocklage (2022), “Marketing insights from text analysis,” *Marketing Letters*, 33 (3), 365-377.

Berger, Jonah and Katherine L. Milkman (2012), “What Makes Online Content Viral?,” *Journal of Marketing Research*, 49 (2), 192-205.

Blei, David M. (2012), “Probabilistic topic models,” *Communications of the ACM*, *55*(4), 77-84.

Boltz, Marilyn G. (2005), “Temporal dimensions of conversational interaction: The role of response latencies and pauses in social impression formation,” *Journal of Language and Social Psychology*, 24 (2), 103-138.

Buller, David B., Beth Al. LePoire, R. Kelly Aune, and Sylvie V. Eloy (1992), “Social perceptions as mediators of the effect of speech rate similarity on compliance,” *Human Communication Research*, 19, 286-286.

Burgoon, Judee K., Laura K. Guerrero, and Valerie Manusov (2011), “Nonverbal signals,” *Handbook of interpersonal communication*, 239-280.

Burgoon, Judee K., and Aaron E. Bacue (2003), “Nonverbal communication skills,” *Handbook of communication and social interaction skills*, 179-219.

Cascio Rizzo, Giovanni Luca, Jonah Berger, Matteo De Angelis, and Rumen Pozharliev (2023), “How Sensory Language Shapes Influencer’s Impact,” *Journal of Consumer Research* (published online March 8), <https://doi.org/10.1093/jcr/ucad017>.

Cesario, Joseph, and E. Tory Higgins (2008), “Making message recipients “feel right” how nonverbal cues can increase persuasion”, *Psychological Science*, 19 (5), 415-420.

Cronin Jr, J. Joseph and Steven A. Taylor (1992), “Measuring service quality: a reexamination and extension,” *Journal of marketing*, *56*(3), 55-68.

Davis, Mark H., Carol Luce, and Stephen J. Kraus (1994), “The heritability of characteristics associated with dispositional empathy,” *Journal of personality*, 62 (3), 369-391.

Davis, Mark H. (1983), “The effects of dispositional empathy on emotional reactions and helping: A multidimensional approach,” *Journal of personality*, 51 (2), 167-184.

Decety, Jean and Philip L. Jackson (2004), “The functional architecture of human empathy,” *Behavioral and cognitive neuroscience reviews*, *3*(2), 71-100.

De Cheveigné, Alain and Hideki Kawahara (2002), “YIN, a fundamental frequency estimator for speech and music,” *The Journal of the Acoustical Society of America*, 111 (4), 1917-1930.

De Jong, Nivja H. and Ton Wempe (2009), “Praat script to detect syllable nuclei and measure speech rate automatically,” *Behavior research methods*, 41 (2), 385-390.

De Ruyter, Ko and Martin G. Wetzels (2000), “The impact of perceived listening behavior in voice-to-voice service encounters,” *Journal of Service Research*, 2 (3), 276-284

Dutton, Jane E., Monica C. Worline, Peter J. Frost, Jacoba Lilius (2006), “Explaining compassion organizing,” *Administrative science quarterly*, 51 (1), 59-96.

Hartmann, Jochen, Heitmann Mark, Schamp Christina, and Netzer Oded (2021) “The Power of Brand Selfies,” *Journal of Marketing Research*, 58 (6), 1159-1177.

He, Jiaxiu, Bingqing Li, and Xin Shane Wang (2023), “Image features and demand in the sharing economy: A study of Airbnb,” *International Journal of Research in Marketing* (published online April 20), https://doi.org/10.1016/j.ijresmar.2023.04.001.

Eisenberg, Nancy and Paul A. Miller (1987), “The relation of empathy to prosocial and related behaviors,” *Psychological bulletin*, 101 (1), 91.

Grewal, Rajdeep, Sachin Gupta, and Rebecca Hamilton (2021), “Marketing insights from multimedia data: text, image, audio, and video,” *Journal of Marketing Research*, 58 (6), 1025-1033.

Grühn, Daniel, Kristine Rebucal, Manfred Diehl, Mark Lumley, and Gisela Labouvie-Vief (2008), “Empathy across the adult lifespan: Longitudinal and experience-sampling findings,” *Emotion*, 8 (6), 753.

Guyer, Joshua J., Leandre R. Fabrigar, and Thomas I. Vaughan-Johnston (2019), “Speech rate, intonation, and pitch: Investigating the bias and cue effects of vocal confidence on persuasion,” *Personality and Social Psychology Bulletin*, 45 (3), 389-405.

Hall, Judith A. and Rachel Schwartz (2019), “Empathy present and future,” *The Journal of social psychology*, 159 (3), 225-243.

Hartmann, Jochen, Mark Heitmann, Christina Schamp, and Oded Netzer (2021), “The power of brand selfies,” *Journal of Marketing Research*, 58 (6), 1159-1177.

Hayes, Andrew F. (2018), “Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation,” *Communication Monographs*, 85 (1), 4-40.

Hama Saeed, Mariwan (2023), “Improved Speech Emotion Classification Using Deep Neural Network,” *Circuits, Systems, and Signal Processing*, 1-20.

Herhausen, Dennis, Lauren Grewal, Krista Hill Cummings, Anne L. Roggeveen, Francisco Villarroel Ordenes, and Dhruv Grewal (2023), “Complaint De-Escalation Strategies on Social Media,” *Journal of Marketing*, 87 (2), 210-231.

Hema, C., and Fausto Pedro Marquez (2023), “Emotional speech Recognition using CNN and Deep learning techniques,” *Applied Acoustics*, 211, 109492.

Hildebrand, Christin, Fotis Efthymiou, Francesc Busquet, William H. Hampton, Donna L. Hoffman, Thomas P. Novak (2020), “Voice analytics in business research: Conceptual foundations, acoustic feature extraction, and applications,” *Journal of Business Research*, 121, 364-374.

Huang, Karen, Michael Yeomans, Alison Wood Brooks, Julia Minson, and Francesca Gino (2017), “It doesn’t hurt to ask: Question-asking increases liking,” *Journal of personality and social psychology*, 113 (3), 430.

Humphreys, Ashlee and Rebecca Jen-Hui Wang (2018), “Automated text analysis for consumer research,” *Journal of Consumer Research*, 44 (6), 1274-1306.

Jacewicz, Ewa, Robert A. Fox, Caitlin O'Neill, and Joseph Salmons (2009), “Articulation rate across dialect, age, and gender,” *Language variation and change*, 21 (2), 233-256.

Juslin, Patrick N. and Petri Laukka (2003), “Communication of emotions in vocal expression and music performance: Different channels, same code?,” *Psychological bulletin*, 129 (5), 770.

Kupor, Daniella M., Zakary L. Tormala, Michael Norton, And Dereck D. Rucker (2014), “Thought calibration: How thinking just the right amount increases one’s influence and appeal,” *Social Psychological and Personality Science*, *5*(3), 263-270.

Lafreniere, Katherine C., Srah G. Moore, and Robert J. Fisher (2022), “The power of profanity: The meaning and impact of swear words in word of mouth,” *Journal of Marketing Research*, *59*(5), 908-925.

Lerner, Jennifer, Ye Li, Piercarlo Valdesolo, and Karim S. Kassam (2015), “Emotion and decision making,” *Annual review of psychology*, 66, 799-823.

Li, Xi, Mengze Shi, and Xin Shane Wang (2019), “Video mining: Measuring visual information using automatic methods,” *International Journal of Research in Marketing*, 36 (2), 216-231.

Li, Yiyi and Ying Xie (2020), “Is a Picture Worth a Thousand Words? An Empirical Study of Image Content and Social Media Engagement,” *Journal of Marketing Research*, 57 (1), 1-19.

Liviatan, Ido, Yacoov Trope, and Nira Liberman (2008), “Interpersonal similarity as a social distance dimension: Implications for perception of others’ actions,” *Journal of experimental social psychology*, 44 (5), 1256-1269.

Luangrath, Andrea W., Joann Peck, and Victor A. Barger (2017), “Textual paralanguage and its implications for marketing communications,” *Journal of Consumer Psychology*, 27 (1), 98-107.

Luangrath, Andrea Webb, Yixiang Xu, and Tong Wang (2022), “Paralanguage Classifier (PARA): An Algorithm for Automatic Coding of Paralinguistic Nonverbal Parts of Speech in Text,” *Journal of Marketing Research*, 60 (2), 388-408.

Morgan, Mallory M. and Jonas Braasch (2022), “Open set classification strategies for long-term environmental field recordings for bird species recognition,” *The Journal of the Acoustical Society of America*, 151 (6), 4028-4038.

McFee, Brian, Colin Raffel, Dawen Liang, Daniel P. Ellis, Matt McVicar, Eric Battenberg, and Oriol Nieto (2015, July), “librosa: Audio and music signal analysis in python,” In *Proceedings of the 14th python in science conference* (Vol. 8, pp. 18-25).

Mikulincer, Mario, Phillip R. Shaver, Omri Gillath, and Rachel A. Nitzberg (2005), “Attachment, caregiving, and altruism: boosting attachment security increases compassion and helping, “*Journal of personality and social psychology*, 89 (5), 817.

Miller, Norman, Geoffrey Maruyama, Rex J. Beaber, and Keith Valone (1976), “Speed of speech and persuasion,” *Journal of personality and social psychology*, 34 (4), 615.

Miller, Joanne L., François Grosjean, and Concetta Lomanto (1984), “Articulation rate and its variability in spontaneous speech: A reanalysis and some implications,” *Phonetica*, 41 (4), 215-225.

Moore, Sarah G. and Katherine C. Lafreniere (2020), “How Online Word-of-Mouth Impacts Receivers,” *Consumer Psychology Review*, 3 (1), 34-59.

Moore, Sarah G. (2012), “Some things are better left unsaid: How word of mouth influences the storyteller,” *Journal of Consumer Research*, 38 (6), 1140-1154.

Moore, Sarah G. (2015), “Attitude predictability and helpfulness in online reviews: The role of explained actions and reactions,” *Journal of Consumer Research*, 42 (1), 30-44.

|  |
| --- |
|  |

Moore, Sarah G. and Brent McFerran (2017), “She said, she said: Differential interpersonal similarities predict unique linguistic mimicry in online word of mouth,” *Journal of the Association for Consumer Research*, 2 (2), 229-245.

Nittrouer, Susan and Joanna H. Lowenstein (2008), “Spectral structure across the syllable specifies final-stop voicing for adults and children alike, “*The Journal of the Acoustical Society of America*, 123 (1), 377-385.

Oleszkiewicz, Anna, Katarzyna Pisanski, Kinga Lachowicz-Tabaczek, and Agnieszka Sorokowska (2017), “Voice-based assessments of trustworthiness, competence, and warmth in blind and sighted adults,” *Psychonomic bulletin & review*, *24*, 856-862.

Packard, Grant, and Jonah Berger (2024). The Emergence and Evolution of Consumer Language Research. *Journal of Consumer Research*, 50 (1).

Packard, Grant and Jonah Berger (2021), “How Concrete Language Shapes Customer Satisfaction,” *Journal of Consumer Research*, 47 (5), 787-806.

Packard, Grant and and Jonah Berger (2017), “How Language Shapes Word of Mouth’s Impact,” *Journal of Marketing Research*, 54 (4), 572-588.

Packard, Grant, Sarah G. Moore, and Brent McFerran (2018), “(I'm) Happy to Help (You): The Impact of Personal Pronoun Use in Customer-Firm Interactions,” *Journal of Marketing Research*, 55 (4), 541-555.

Packard, Grand, Jonah Berger, and Reihane Boghrati (2023), “How Verb Tense Shapes Persuasion. *Journal of Consumer Research* (published online January 23), https://doi.org/10.1093/jcr/ucad006.

Paetzold, Gustavo and Lucia Specia (2016, June), “Inferring psycholinguistic properties of words,” In *Proceedings of the 2016 conference of the north american chapter of the association for computational linguistics: Human language technologies* (pp. 435-440).

Patrick, Vanessa M., and Henrik Hagtvedt (2012a), “I don’t” versus “I can’t”: When empowered refusal motivates goal-directed behavior,” *Journal of Consumer Research*, 39 (2), 371-381

Patrick, Vanessa M., and Henrik Hagtvedt (2012b), “How to say “No”: Conviction and identity attributions in persuasive refusal,” *International Journal of Research in Marketing*, 29 (4), 390-394.

Parasuraman, Arun, Leonard L. Berry, and Valarie A. Zeithaml (1991), “Understanding customer expectations of service,” *MIT sloan management review*.

Pennebaker, James W., Ryan L. Boyd, Kayla Jordan, and Kate Blackburn (2015), “*The development and psychometric properties of LIWC2015”*.

Pogacar, Ruth, L. J. Shrum, and Tina M. Lowrey (2018), “The Effects of Linguistc Devices on Consumer Information Processing and Persuasion: A Language Complexity x Processing Mode Framework,” *Journal of Consumer Psychology*, 28 (4), 689-711.

Pogacar, Ruth, Alican Mecit, Fei Gao, L. J. Shrum, and Tina M. Lowrey (2022), “Language and consumer psychology.”

Preston, Stephanie D. and Frans BM De Waal. (2002), “Empathy: Its ultimate and proximate bases,” *Behavioral and brain sciences*, 25 (1), 1-20.

Quené, Hugo (2008), “Multilevel modeling of between-speaker and within-speaker variation in spontaneous speech tempo,” *The Journal of the Acoustical Society of America*, 123 (2), 1104-1113.

Ray, George B. (1986), “Vocally cued personality prototypes: An implicit personality theory approach,” *Communications Monographs*, 53 (3), 266-276.

Robin, Olivier and Catherine Plante (2022), “An illustrated tutorial for logarithmic scales and decibels in acoustics,” *The Journal of the Acoustical Society of America*, 152 (5), 2880-2892.

Rochman, Daniel, Gary Diamond, and Ofer Amir (2008), “Unresolved anger and sadness: Identifying vocal acoustical correlates,” *Journal of Counseling Psychology*, 55 (4), 505.

Rodero, Emma (2020), “Do Your Ads Talk Too Fast To Your Audio Audience?: How Speech Rates of Audio Commercials Influence Cognitive and Physiological Outcomes,” *Journal of advertising research*, 60 (3), 337-349.

Schroeder, Juliana and Nicholas Epley (2015), “The sound of intellect: Speech reveals a thoughtful mind, increasing a job candidate’s appeal,” *Psychological science*, 26 (6), 877-891.

Sela, Aner, S. Christian Wheeler, and Gülen Sarial-Abi (2012), “We are not the same as you and I: Causal effects of minor language variations on consumers' attitudes toward brands,” *Journal of Consumer Research*, 39 (3), 644-661.

Singh, Jagdip and Deepak Sirdeshmukh (2000), “Agency and trust mechanisms in consumer satisfaction and loyalty judgments,” *Journal of the Academy of marketing Science*, *28*, 150-167.

Smith, Vicki L. and Herbert H. Clark (1993), “On the course of answering questions. *Journal of memory and language*, 32 (1), 25-38.

Šošić, Milena and Jelena Graovac (2022), “Effective methods for email classification: Is it a business or personal email?,” *Computer Science and Information Systems*, 19 (3), 1155-1175

|  |
| --- |
|  |

Street Jr, Richard L. (1982), “Evaluation of noncontent speech accommodation,” *Language & Communication*, 2 (1), 13-31.

Street Jr, Richard L. (1984), “Speech convergence and speech evaluation in fact-finding interviews,” *Human Communication Research*, 11 (2), 139-169.

Tibshirani, Robert (1996), “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58 (1), 267-288.

Van Zant, Alexander B. and Jonah Berger (2020), “How the voice persuades,” *Journal of Personality and Social Psychology*, 118 (4), 661-682.

Van Zant, Alex, Jonah Berger, Grant Packard, and Harry Wang (2023), “The Power of Pausing,” *Under Revision*.

Wang, Xin Shane, Shijie Lu, X. I. Li, Mansur Khamitov, and Neil Bendle (2021), “Audio Mining: The Role of Vocal Tone in Persuasion,” *Journal of Consumer Research.*

Wang, Xin Shane, Jiaxiu He, David J. Curry, and Jun Hyun Ryoo (2022), “Attribute embedding: Learning hierarchical representations of product attributes from consumer reviews,” *Journal of Marketing*, 86 (6), 155-175.

Wang, Ze, Aaron D. Arndt, Surendra Singh, Monica Biernat, and Fan Liu (2013). “You Lost Me at Hello: How and when accent-based biases are expressed and suppressed,” *International Journal of Research in Marketing*, 30 (2), 185-196.

**Web Appendix: How Speaking Rate Shapes Consumer Response**

**Study 1**

Figure WA1: Model Results

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **(1)** | **(2)** | **(3)** | **(4)** |
| **Slower Artic. Rate**  Controls  *Call*  Call Reasons  Account  Order  Product  Return  Shipping  Call Topics  Call Severity  Solved Issue  Call Duration  *Employee* | **.182\*\* (.080)**  -1.617\*\* (.855)  -.344\*\* (.762)  -.849\*\* (.850)  -.873\*\* (.762)  -.873\*\* (.762)  Included  .073\*\* (.087)  .431\*\* (.215)  -.540\*\* (.222) | **.191\*\* (.077)**  -1.459\*\* (.858)  .073\*\* (.776)  -.232\*\* (.872)  -.479\*\* (.768)  -.437\*\* (.774)  Included  .103\*\* (.090)  .365\*\* (.213)  -.636\*\* (.223) | **1.923\*\* (.421)**  .008\*\* (.015)  .270\*\* (.437)  .053\*\* (.102)  .044\*\* (.073)  .074\*\* (.123)  Included  1.383\*\* (.309)  2.437\*\* (1.262)  .163\*\* (.096) | **.135\*\* (.067)**  -.986\*\* (.386)  .265\*\* (.183)  -.129\*\* (.163)  Included  .344\*\* (.177) |
| Tenure | .002\*\* (.082) | -.024\*\* (.083) | 1.043\*\* (.234) |  |
| Female | -.017\*\* (.219) | -.027\*\* (.223) | .805\*\* (.458) |  |
| Pauses  Pitch  Concreteness  Questions  Psych. Processes  Affect  Bio  Cognition  Drives  Informal  Percept  Relative  Social  Time | .221\*\* (.159)  -.062\*\* (.108)  .200\*\* (.117)  .241\*\* (.108)  .102\*\* (.102)  -.083\*\* (.082)  -.151\*\* (.089)  .044\*\* (.081)  -.240\*\* (.103)  -.032\*\* (.091)  -.116\*\* (.109)  .056\*\* (.103)  -.103\*\* (.099) | .276\*\* (.160)  -.024\*\* (.107)  .281\*\* (.167)  .259\*\* (.108)  .177\*\* (.114)  -.102\*\* (.085)  -.076\*\* (.093)  .014\*\* (.082)  -.219\*\* (.108)  -.082\*\* (.097)  -.149\*\* (.112)  .045\*\* (.106)  -.100\*\* (.099) | 2.726\*\* (1.213)  .753\*\* (.208)  3.010\*\* (1.284)  1.824\*\* (.534)  1.458\*\* (.368)  .807\*\* (.163)  .530\*\* (.129)  1.105\*\* (.212)  .429\*\* (.124)  .889\*\* (.205)  .757\*\* (.203)  1.162\*\* (.294)  .715\*\* (.174) | .164\*\* (.068)  .143\*\* (.076)  .114\*\* (.083)  -.133\*\* (.067)  -.161\*\* (.080)  .061\*\* (.071) |
| *Customer*  Tenure  Female  Age  18-24  25-34  35-44  55+  Region  East  Midwest  South  West  Artic. Rate  Pauses  Pitch  Concreteness  Questions  Psych. Processes  Affect  Bio  Cognition  Drives  Informal  Percept  Relative  Social  Time  R-squared | -.033\*\* (.077)  -.027\*\* (.201)  .184\*\* (.328)  .085\*\* (.275)  .366\*\* (.305)  .231\*\* (.423)  .532\*\* (.285)  .189\*\* (.323)  .601\*\* (.316)  .104\*\* (.291)  .159\*\* (.089)  .615\*\* (.159)  -.024\*\* (.093)  -.071\*\* (.134)  -.162\*\* (.092)  .102\*\* (.132)  -.052\*\* (.093)  .056\*\* (.092)  -.006\*\* (.093)  .013\*\* (.132)  .230\*\* (.087)  .108\*\* (.086)  .028\*\* (.093)  .049\*\* (.089)  .600 | -.026\*\* (.079)  -.003\*\* (.205)  .114\*\* (.335)  .009\*\* (.279)  .386\*\* (.308)  .006\*\* (.432)  .574\*\* (.286)  .247\*\* (.328)  .723\*\* (.320)  .064\*\* (.293)  .151\*\* (.090)  .597\*\* (.163)  -.044\*\* (.096)  -.080\*\* (.148)  -.115\*\* (.097)  .128\*\* (.135)  .008\*\* (.095)  .080\*\* (.102)  .024\*\* (.097)  .037\*\* (.139)  .242\*\* (.092)  .122\*\* (.091)  .013\*\* (.094)  .056\*\* (.089)  .604 | .844\*\* (.176)  .698\*\* (.347)  3.168\*\* (2.734)  2.384\*\* (1.802)  3.653\*\* (2.977)  1.353\*\* (1.441)  9.713\*\* (7.923)  4.486\*\* (4.080)  7.179\*\* (6.034)  2.406\*\* (1.929)  1.556\*\* (.344)  9.905\*\* (4.394)  .885\*\* (.207)  .847\*\* (.304)  .520\*\* (.130)  1.192\*\* (.393)  .783\*\* (.175)  1.289\*\* (.312)  .959\*\* (.228)  .928\*\* (.320)  1.988\*\* (.465)  1.370\*\* (.288)  1.017\*\* (.229)  1.121\*\* (.249) | -.116\*\* (.162)  -.137\*\* (.139)  .134\*\* (.192)  .384\*\* (.175)  .496\*\* (.220)  -.026\*\* (.181)  .164\*\* (.067)  .333\*\* (.092)  -.190\*\* (.074)  .103\*\* (.076)  .175\*\* (.065)  .546 |
| N | 189 | 189 | 189 | 189 |

*Notes*: \**p* < .1, \*\**p* < .05. Standard errors are in parentheses. Giftcard reason, Age 45-54, and Other region are baselines. LDA topic coefficients are all non-significant at *p* > .01. **Column 1:** Main model, **Column 2**: Alternative measure of articulation rate, **Column 3**: Ordinal logistic regression (estimates are ORs), **Column 4**: Lasso Penalized regression.

**Study 2**

*Exclusion and Demographic Information*. Participants (*N* = 500) were recruited through Prolific. People (*N* = 44) were excluded if they experienced audio issues or if they failed at least one of the two attention checks asking them “what was the issue about? giftcard, internet connection, bills”, and “the agent was? male, female”. The final sample consists of 456 (50.5% women, 48% men, 1.5% other; mean age = 41.1 years).

*Study Design*. People were exposed to a scenario that read, “Imagine that you’ve had issues with various mobile phone service providers. Recently, however, you signed up for a new contract. Not only are the charges fair, but it seems like the problems you've faced are finally over! Everything seems promising until you receive the first bill, only to discover that you’ve been charged much more than expected. You’re quite angry, so you call customer service to try to fix things.”

Then, they listened to the agent’s response through one of two professional voice recordings. Audio recordings are accessible at https://osf.io/5kt7h. The script reads, “*Umh I see. Let me look into your account and billing details. So it looks like there was an error in our billing system; that’s why your charges were higher. We’ve got a couple of potential solutions. One is to provide a credit for the excess amount charged. Another option is to offer a discounted rate on your future bills.* *Which of these two options do you prefer?*”

*Manipulation Check.* Participants in the slow articulation rate condition perceived the agent’s speaking speed to be slower (*M*slow = 3.78) than those in the fast articulation rate condition (*M*fast = 5.34, *F*(1, 454) = 287.23, *p* < .001, η2 = .387).

**Study 3**

*Exclusion and Demographic Information*. Participants (*N* = 400) were recruited through Prolific. People (*N* = 38) were excluded if they experienced audio issues or if they failed at least one of the two attention checks asking them “which was the issue? allergies, headaches, digestive, skin conditions”, and “the doctor was? male, female”. The final sample consists of 362 (54.6% women, 44.3% men, 1.1% other; mean age = 43.1 years).

*Study Design*. People were exposed to a scenario that read, “Imagine that you have been experiencing persistent migraines, causing you significant discomfort. They make it challenging for you to concentrate, work, and enjoy all the activities you once loved. After considering the options, you decide to visit a doctor in the hopes of finding relief. Unfortunately, however, the doctor is not very understanding or helpful, and didn’t seem that concerned about your wellbeing.  So you decide to visit a second doctor.”

Then, they listened to the agent’s response through one of two professional voice recordings. Audio recordings are accessible at https://osf.io/5kt7h. The script reads, “*I see. Have you ever given any previous treatments or medications a shot for your migraines? I'm asking because it helps me get a better understanding of what has and hasn't worked in the past. Also, need to hear more about any specific triggers or patterns you've noticed when it comes to your migraines. It could be certain foods that act as troublemakers, stress that tends to amp them up, sleep patterns or even environmental factors. It's all about knowing what to avoid or manage to reduce those migraines you know”.*

*Manipulation Check.* Participants in the slow articulation rate condition perceived the agent’s speaking speed to be slower (*M*slow = 3.78) than those in the fast articulation rate condition (*M*fast = 4.64, *F*(1, 359) = 69.12, *p* < .001, η2 = .161).

**Study 4**

*Exclusion and Demographic Information*. Following the preregistration, 846 participants were recruited through Prolific. To account for failed attention checks and audio issues, we collected about 5% more responses than preregistered. People (*N* = 44) were excluded if they experienced audio issues or if they failed at least one of the two attention checks of Study 3. The final sample consists of 802 (54.4% women, 43.1% men, 2.5% other; mean age = 40.5 years).

*Study Design*. Participants in the baseline (i.e., high need for empathy) condition were exposed to the scenario of Study 3. Participants in the low need for empathy condition were exposed to a scenario that read, “Imagine that you have been experiencing a mild headache once in a while. You’re pretty sure it's no big deal, but to make sure, you decide to call your trusted and caring doctor. They immediately tell you not to worry and that the problem can be easily solved.” The audio recordings were the same of Study 3.

*Manipulation Checks.* Participants in the slow articulation rate condition perceived the agent’s speaking speed to be slower (*M*slow = 3.88) than those in the fast articulation rate condition (*M*fast = 4.52, *F*(1, 800) = 79.06, *p* < .001, η2 = .090). Furthermore, people in the high need for empathy condition (*M*high need = 5.21) looked for doctor’s empathy more than those in the low need for empathy condition (*M*low need = 4.89, *F*(1, 800) = 12.36, *p* < .001, η2 = .015).

*Satisfaction with the Interaction*. Effects were similar for satisfaction. Main effects of both articulation rate (*F*(1, 798) = 3.99, *p* = .046, η2 = .005), and need for empathy (*F*(1, 798) = 9.56, *p* = .002, η2 = .012), were qualified by the predicted articulation rate × need for empathy interaction (*F*(1, 798) = 3.47, *p* = .063, ηp2 = .004). Consistent with Study 3, in the baseline condition, speaking slower increased satisfaction with the interaction (*M*slow = 6.00 vs. *M*average= 5.67, *F*(1, 798) = 7.41, *p* = .007, η2 = .021). Consistent with the hypothesized underlying role of empathy, however, when people’s need for empathy was lower, the beneficial effect of speaking slowly was mitigated (*M*slow = 5.57 vs. *M*average= 5.56, *F*(1, 798) = .01, *p* = .925, η2 < .001).

*Evaluation*. Effects were similar for evaluations. A main effect of need for empathy (*F*(1, 798) = 8.28, *p* = .004, η2 = .010) was qualified by the predicted articulation rate × need for empathy interaction (*F*(1, 798) = 4.70, *p* = .030, ηp2 = .006). Consistent with Study 3, in the baseline condition, speaking slower made the doctor seem like a better doctor (*M*slow = 6.10 vs. *M*average= 5.86, *F*(1, 798) = 4.79, *p* = .029, η2 = .014). Consistent with the hypothesized underlying role of empathy, however, when participants’ need for empathy was lower, the beneficial effect of speaking slowly was mitigated (*M*slow = 5.71 vs. *M*average= 5.80, *F*(1, 798) = .76, *p* = .382, η2 =. 002).

1. When the human ear perceives sound, it receives sound waves that vary in loudness and pitch. However, the way the human ear perceives loudness is not linearly related to the actual amplitude of the sound wave. Instead, it perceives loudness on a logarithmic scale, meaning that doubling the amplitude of the sound wave, for example, will not results in a doubling of perceived loudness (Robin and Plante 2022). [↑](#footnote-ref-1)
2. Two research assistants (blinded to hypotheses) rated a random sample of 30 calls on how quickly the person spoke (1 = not at all, 7 = very much; *r* = .60). [↑](#footnote-ref-2)
3. Employee articulation rate and call duration are not correlated (*r* = –.02, *p* = .774). [↑](#footnote-ref-3)
4. One might wonder whether employees who speak more slowly also tend to pause more frequently. Employee articulation rate and pausing, however, are not correlated (*r* = .04, *p* = .575). [↑](#footnote-ref-4)
5. We do not control for loudness due to its potential susceptibility to recording-specific volume characteristics (e.g., some recording audios are louder at the outset) rather than reflecting individual vocal loudness. [↑](#footnote-ref-5)
6. While one might wonder whether speaking too slowly backfires, note that introducing the articulation rate’s quadratic term finds no significance (*b* = .003; *SE* = .057, *t* = .06, *p* = .956). [↑](#footnote-ref-6)
7. Note that rather than signaling thoughtfulness, some work suggests it can lead people to be perceived negatively (e.g., more uncertain or incompetent, Boltz 2005; Street 1982; 1984), [↑](#footnote-ref-7)
8. Speaking more quickly was also not linked to customers using words or phrases indicating difficulties understanding. Following Humphrey and Wang (2018), we used a bottom-up approach (guided by the most frequent words in our sample) to create a lexicon of phrases signaling customer understanding (i.e., “I understand,”, “I see,” “I hear you,” “okay,” “ok,” “makes sense,” and “got it”) and measured the level of understandability in customers language at the turn level. Articulation rate and customer understandability were unrelated (*r* = –.02) casting further doubt on the notion that speaking more quickly made things difficult to understand. [↑](#footnote-ref-8)
9. As discussed in Study 1, while articulation rate is the speed at which people speak, response time is the speed at which someone *responds* to others (i.e., pause time between conversational exchanges). [↑](#footnote-ref-9)
10. Unlike prior work that found a negative effect of speaking speed, in this case participants were imagining a social interaction, and the recording was specified as a *response* to something the participants said. Thus, inferences about empathy should be more likely. [↑](#footnote-ref-10)
11. Participants (*N* = 60) randomly listened to one of the two stimuli and rated “How fast did the agent speak?” (1= very slowly, 7 = very fast). [↑](#footnote-ref-11)
12. As in Study 2, there was no difference between conditions on pitch (53 vs. 52), intonation (.16 vs. .16), loudness (.03 vs. .03), or pausing (0 vs. 0). [↑](#footnote-ref-12)
13. Pretest participants were asked “how much empathy were you looking for from the doctor?” (1 = very little empathy, 7 = a lot of empathy). [↑](#footnote-ref-13)