OXFORD UNIVERSITY PRESS | Journal of Consumer Research

Please see our manuscript evaluation criteria and reviewer guidelines.

When Language Matters

Journal:	*Journal of Consumer Research*
Manuscript ID	JCONRES-2022-102.R4
Manuscript Type:	Article
Article Type:	(Multi-)methods/empirical quant paper (Uses datasets, quantitative models or multiple methods to study consumer behavior
Basic Disciplines:	Linguistics and Semiology, Psychology
Theoretical Topics:	Social Influence and Norms, Persuasion
Substantive Topics:	Consumer Experience, Consumer Satisfaction, Marketing Communications, Services
Methods:	Computational Linguistics, Econometrics, Experimental Design/Analysis, Multivariate Data Analysis

When Language Matters

ABSTRACT

Text analysis is increasingly used for consumer and marketing insight. But while work has shed light on *what* firms should say to customers, *when* to say those things (e.g., within an advertisement or sales interaction) is less clear. Service employees, for example, could adopt a certain speaking style at a conversation's start, end, or throughout. *When* might specific language features be beneficial? This paper introduces a novel approach to address this question. To demonstrate its potential, we apply it to warm and competent language. Prior research suggests an affective (i.e., warm) speaking approach leads customers to think employees are less competent, so a cognitive (competent) style should be prioritized. In contrast, our theorizing, analysis of hundreds of real service conversations from two firms across thousands of conversational moments (N = 23,958), and four experiments (total N = 1,589) offer a more nuanced perspective. Customers are more satisfied when employees use *both* cognitive and affective language, but at separate, specific times. Ancillary analyses show how this method can be applied to other language features. Taken together, this work offers a method to explore *when* language matters, sheds new light on the warmth/competence trade-off, and highlights ways to improve the customer experience.

Keywords: Language, Communication, Dynamics, Warmth and Competence, Customer Service.

Language is an integral part of communication. Advertising copy shapes purchase, service language shapes customer retention, and the words in word of mouth shape consumer behavior (e.g., McGuire 2000; Ordenes et al. 2014; Pogacar, Shrum, and Lowrey 2018; Schellekens, Verlegh, and Smidts 2010). Consistent with language's importance, decades of research has considered how employees should speak to customers (e.g., Parasuraman, Zeithaml, and Berry 1985; Blanding 1989) and natural language processing tools are shedding new light on language that increases communication's impact (Berger et al. 2020; Humphreys and Wang 2018).

But while it's clear that *what* companies, employees, and consumers say matters, might *when* they say it within a given communication also play an important role?

Calling customer service, for example, or speaking with a salesperson usually involves a conversation. Customers say something, employees respond, and the two go back and forth. While research suggests that asking questions, using first person pronouns, or speaking in a rational, competence-oriented way can improve customer satisfaction (Drollinger, Comer, and Warrington 2006; Marinova, Singh, and Singh 2018; Packard, Moore, and McFerran 2018), should employees do these things throughout an interaction? Or might doing so at certain conversational points be more beneficial?

Take greetings. Call center agents could say "Who do I have the pleasure of speaking with?" or "How may I assist you?" Both are common openings, but the first is warmer while the latter focuses on competence. The same goes for conversation endings such as "It was my pleasure. Take care now" or "I'm glad I could solve that for you. Bye now." The former uses warmer, more affective language and the latter a more cognitive, competence-oriented approach. While a great deal of research suggests prioritizing competence in consumer communications (e.g., Gunturkun, Haumann, and Mikolon 2020; Kirmani et al. 2017; Li, Chan, and Kim 2019; Marinova et al. 2018), is that actually the best course of action in these conversational moments?

This paper moves beyond asking *whether* particular language features matter, to introducing an approach for studying *when*. Conversations are a key part of social interaction (Huang et al. 2017), but the moment-to-moment content variation in conversations makes them difficult to analyze (Reece et al. 2022; Zhang, Wang, and Chen 2020). To address these challenges, we use functional data analysis (FDA; e.g., Foutz and Jank 2010), recovering time-based sensitivity trajectories and documenting the dynamic relationship between language and important marketing outcomes.

To demonstrate the approach, and its potential, we apply it to language linked to the two central dimensions of person perception — warmth and competence (Fiske, Cuddy, and Glick 2007). A multi-method investigation, including analysis of thousands of moments across hundreds of service conversations at two firms, and four experiments, suggests customers are more satisfied (and spend more) when employees use *both* cognitive and affective language, but at separate, specific times. Ancillary analyses apply our approach to other language features.

This paper makes three main contributions. First, most narrowly, we deepen insight into the so-called warmth/competence trade-off. While research suggests emphasizing only one of these in a given interaction (i.e., prioritize warmth or competence but not both; Dubois, Rucker and Galinsky 2016; Godfrey, Jones, and Lord 1986; Fiske et al. 2007; Holoien and Fiske 2013), we find this "trade off" may not be so stark. Instead, results reveal that service employees should prioritize *both* cognitive and affective language, but at different points in time. Each is beneficial (or costly) at different, specific moments within an interaction.

Second, we demonstrate that understanding when different language features matter can improve marketing outcomes. While one might wonder whether employees are already sufficiently warm at the start and end, for example, two field data sets suggest this is not the case. Results reveal that employees may benefit from using warmer language than they currently do at the start of interactions. Ancillary analyses reveal *when* other language features recommended by prior research (e.g., question asking and first-person pronouns) matter as well. Our approach can help improve customer service, aid employee assessment and development, and fine-tune artificial intelligence (AI) chatbots' effectiveness. It can also be used to shed light on word of mouth, sales interactions, and marketing communications more broadly.

Third, we introduce a novel modeling approach using functional data analysis and Group Lasso to tackle the high dimensionality, irregularity, and sparsity inherent in conversational data. An emerging stream of work has begun to study conversations (Ordenes and Grewal 2017; Yeomans, Schweitzer, and Brooks 2022) and advertising, word of mouth, and other marketing interactions involving conversational language. Across these and other contexts, our method can help researchers better understand not only what language matters, but *when*. This approach provides a framework for understanding language dynamics, and their impact, within consumer research, and beyond. To help other researchers leverage this approach, we created a free user-friendly web application.[1]

TALKING TO CUSTOMERS

[1] Non-technical users can upload a text file and perform dynamic "when" analysis on their own datasets without the use of programming language at whenlanguagematters.net. Customizable R code is also available at the same website.

Talking to customers is important. Companies spend over a trillion dollars a year on sales and service alone, making it the single largest strategic investment for most firms, and nearly tripling what they spend on other marketing communications (Cespedes and Wallace 2017; Morgan 2017). Further, these costs are rising as channel complexity and technology make it harder to deliver great service (McBain 2020).

Consistent with its importance, a great deal of research has tried to understand and improve these interactions. Thousands of articles have studied service quality (Parasuraman and Zeithaml 2002; Snyder et al. 2016), examining how consumers evaluate salespeople (e.g., Zeithaml, Berry, and Parasuraman 1996), service initiatives shape customer attitudes (e.g., Bolton and Drew 1991), and service quality impacts firms (Rust and Chung 2006).

Along these lines, research has explored the role of language in marketing communications, sales, and service (cf. Pogacar et al. 2022 for a recent review). Experienced salespeople are more likely to use questions like "Could you tell me more?" (Castleberry, Shepherd, and Ridnour 1999), for example, and asking such questions can signal attention and empathy, fostering effective conversations (Brody 1994; Brooks and John 2018; Drollinger and Comer 1997). Similarly, concrete language (e.g., "jeans" instead of "clothes") encourages purchase because it suggests service agents are listening (Packard and Berger 2021) and first-person singular ("I") pronouns enhance customer satisfaction because it makes employees seem more agentic and empathetic (Packard et al. 2018).

But while a growing body of research demonstrates language's importance, less is known about *when* particular language features should be used. Should such language features be used throughout a conversation, for example, or might they be more beneficial at certain moments? And might they backfire in others?

WHEN LANGUAGE MATTERS

To illustrate the value of *when*, we examine the "warmth/competence trade-off" (Durante, Tablante, and Fiske 2017). Warmth and competence are central dimensions of social cognition, accounting for almost all person perception (Fiske et al. 2007). Warmth captures affective expression and attention to emotions while competence focuses on agency, rationality, and cognitive efficiency (Abele and Wojciszke 2007). Above all else, people evaluate others on these fundamental dimensions (Judd et al. 2005).

Importantly, however, a great deal of research suggests these two dimensions are inversely related. Being affectively engaged makes people seem less competent, while being cognitively-oriented makes people seem less warm (Fiske et al. 2007). This has led researchers to suggest people should try to be warm or competent, but not both (Dubois et al. 2016; Godfrey et al. 1986; Fiske et al. 2007; Holoien and Fiske 2013; Wojciszke et al. 1998).

Many marketing researchers have suggested a competence-oriented approach is best (e.g., Kirmani et al. 2017; Li et al. 2019; cf. review in Gunturkun et al. 2020). Solution-oriented service advisors reportedly enhance customer satisfaction more than socially-oriented agents (van Dolen et al. 2007) and service employees who use emoticons are seen as warmer, but less competent, leaving customers less satisfied (Li et al. 2019). Competence is said to be prized over warmth in service interactions (Kirmani et al. 2017) because consumers are goal-oriented and can't achieve their goals if a service provider isn't sufficiently skilled (Kirmani and Campbell 2004). Even research proposing a "golden quadrant" in which marketers might benefit from *both* warmth and competence ultimately suggested that only competence drove positive outcomes (Aaker, Garbinsky, and Vohs 2012).

Particularly relevant to the current investigation, Marinova, Singh, and Singh (2018) concluded that employee affective language hindered the benefit of a more cognitive, solution-oriented speaking style, both overall and when examined within three interaction phases. Similarly, Singh, Marinova, Singh and Evans' (2018) modeling of agent language in insurance sales found that warm language curtailed or even neutralized the benefits of more cognitive, solving-oriented language.

Indeed, when engaging customers, firms tend to prioritize competent problem solving rather than relational warmth (Dixon, Freeman, and Toman 2010; Jasmand, Blazevic, and de Ruyter 2012). When we asked 160 customer service managers and workers about the most important service priority, 80.8% indicated "competently addressing the customer's needs" (vs. "warmly relating to the customer"), and 76.1% indicated their company training prioritizes competence. Only 21.3% indicated their firm trains employees to be both competent and warm.

But should service agents necessarily prioritize a competence-oriented, cognitive manner of speaking throughout an interaction? And how does this fit with older work encouraging employees to speak affectively to show customers they care (e.g., de Ruyter and Wetzels 2000; Parasuraman et al. 1985; Spiro and Weitz 1990)?

THE CURRENT RESEARCH

We propose that, rather than asking whether employees should speak cognitively or affectively, it is important to consider conversational moments. Rather than only considering whether one type of language is better than the other overall, we suggest that a more granular, turn-by-turn analysis will show that what language is effective depends on *when* in a conversation it occurs.

Research on conversational analysis and implicature supports this suggestion. Each turn contributes to a conversation's ultimate meaning and outcome (Goffman 1981; Schegloff 1999). A conversational dialogue that "works" is one in which each meaningful statement is satisfied by a relevant and meaningful response (Grice 1991). Indeed, Grice's famous conversational principles (e.g., relation and manner) are explicitly conceptualized as localized, turn-by-turn exchanges rather than at an aggregate level.

Building on this work, we suggest that a given language feature's importance should be moderated by conversational moment. Early in service interactions, we suggest affective language will be more effective than task-oriented, cognitive language. While the norms of conversational openings demand a sequence of pleasantries (Schegloff 1999), these turns can vary in the extent to which they focus on warmth or competence. Agents could start with more cognitive, competence-oriented language (e.g., "How may I assist you?") or more affective, warm language (e.g., "How are you today?"). Social norms suggest warm behaviors such as relationship-building, empathy, or apology can be useful before turning to the speakers' specific goals (Clark et al. 2013; Gabor 2011; Kaski, Niemi, and Pullins 2018; Radu et al. 2019). Consequently, while "How may I assist you?" is a common opening, it jumps straight into problem solving rather than establishing a warm, relational base (Placencia 2004), which should make it less effective in early conversational moments.

But while starting with more affective language may be important, it should only go so far. Eventually employees must competently address the customer's goals and needs. Conversation analysis notes the importance of shifting discourse from greetings and preliminaries to "getting down to business" (Bolden 2008; Pallotti and Varcasia 2008). In conversational turns central to the "business" of customer service, for example, employees may be better off using language like “I’m going to resolve this” rather than a warmer “I’m happy to help with this.” Consequently, a more analytic, cognitive communication style should be beneficial in conversation’s middle moments.

Finally, more affective language may be beneficial at a conversation’s close. Closing with more rational, cognitive language may seem like “dismissals” (Frank 1982). Consistent with our suggestion, wrapping up an interaction in a considerate or empathetic manner is thought to be a key feature of successful conversations (Schegloff and Sacks 1973), and may help align participants’ conceptions of the interaction (Aston 1995).

Again, we are not just suggesting it is good to be polite and positive at the beginning and end of conversations. Instead, we propose prioritizing different *kinds* of language at such conversational moments. Both “My pleasure. Take care now” and “I’m glad we could solve that for you. Bye now” signal the conversation’s end in a polite and positive way. But because the former involves warmer, more affective language, we suggest it will be more beneficial.

To test these predictions, we analyze linguistic (verbal) features over conversational time to examine *when* employee language has a positive, null, or negative relationship with customer satisfaction. A multimethod approach, including two field data sets and four experiments, tests this perspective. To examine these relationships in the field, we devise a novel empirical approach and analyze two large turn-level data sets of customer service conversations from companies in different market sectors. To assess our approach’s contribution, we compare it to (a) traditional static approaches, (b) simpler, more discrete (rather than continuous) dynamics considered in prior literature, and (c) other simplified or restricted models. We demonstrate its robustness not only for customer satisfaction, but also purchase behavior and willingness to recommend. Four experiments then (a) directly test causality and validity of the model results,

(b) consider alternative dynamics, and (c) explore robustness across various naturalistic and carefully controlled stimuli (Studies 3, 4A, 4B, and 5).

Finally, we demonstrate how our approach can offer new insight into other language features and discuss its potential (and limitations) for understanding and optimizing communication more broadly.

STUDY 1: RETAILER FIELD DATA

To provide an initial test of our theorizing, we collected a random sample of 200 customer service calls from a large online retailer. A professional transcription company converted the recordings to text, separating each conversational turn (e.g., turn 1 (agent): "How can I help you?", turn 2 (customer): "I can't find …"). Part of the conversation was inaudible for fifteen recordings provided, leaving 12,410 turns from 185 conversations (handled by a total of 130 agents).[2] The average conversation lasted 6.19 minutes (SD = 3.97) and included 66.75 turns (SD = 44.49). See Web Appendix A for additional conversation descriptive statistics.

Independent Measures: Agent Affective and Cognitive Language

Following prior work (Berry et al. 1997; Marinova et al. 2018; Singh et al. 2018), we measure affective and cognitive language through Linguistic Inquiry and Word Count's (LIWC; Pennebaker et al. 2015) affective processes module (i.e., 1,388 validated words and word stems

[2] While the number of conversations analyzed may seem smaller than contexts like online reviews, it is quite large when it comes to the dynamics of marketing conversations (see Web Appendix Table A1). This is in part because the unit of analysis in such research entails modeling a time series of units within each conversation, described as slices, stages, segments, or turns (e.g., Marinova et al. 2018; Singh et al. 2018; Singh et al. 2020).

such as happy and horrible). Warmth is conveyed through emotional expression. Using affective words like *happy* (e.g., "I'm happy you like the pants") or *horrible* ("That's horrible") signals that an employee is attending to a customer's emotional state or expressing their own.

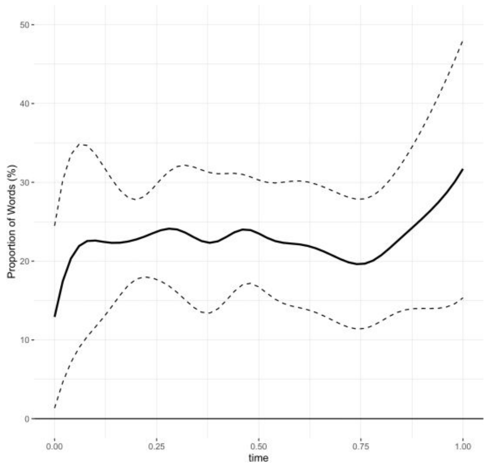
Cognitive language involves rational expression suggesting instrumentality, intelligence, and agency. Using cognitive words like *diagnose* (e.g., "Let's diagnose the cause") or *think* ("I think that will do it") signals that an agent is cognitively working to address the customer's needs. Following prior work, cognitive language is measured through LIWC's cognitive processes module, which contains 780 relevant words and word stems (e.g., diagnose or think).

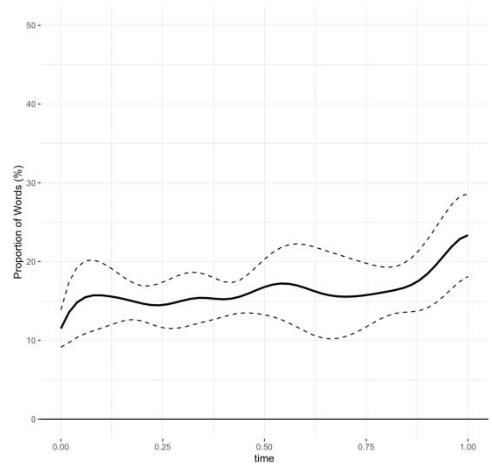
Figure 1 illustrates what agents do currently (i.e., their average affective and cognitive language over the course of conversations). Affective language, for example, makes up roughly 13-24% of words in opening turns. Notably, while conversations often start with pleasantries or greetings, affective language is not particularly high at the outset, indicating that agents do not use especially warm language at this time. Similarly, agent use of cognitive language does not peak in the middle, "business" portion of the conversation where we suggest it may be important. Finally, as indicated by the 95% confidence dotted lines, there is considerable variation across agents in the language used over the course of conversation.[3]

Figure 1: Focal Features over Conversational Time

(A) Agent Affective Language (B) Agent Cognitive Language

[3]The ratio of the two language types over time (Web Appendix A Figure A1) also suggests that agents do not prioritize warm, affective language over competence oriented, cognitive language at the start or end of conversations.





Note: The y-axis depicts conversational turn-level measurement of a focal language feature across non-zero turns (i.e., the percentage of words in a turn corresponding to affective and cognitive language respectively).

Dependent Measure

Study 1 focuses on perceived helpfulness, a key measure of customer satisfaction (Cronin and Taylor 1992; Parasuraman Berry, and Zeithaml 1991). We collected the firm's measure of this for each call (1= not at all helpful, 4= very helpful, measured at the end of the call). For robustness we also later consider a behavioral measure—the number of purchases made in the 30 days following the call.

Controls

While our interest is in warm and competent language, one could wonder whether any relationship between these features and customer satisfaction is driven by other observable factors. Consequently, we control for a range of control variables pertaining to the call, agent, or customer that are conceptually or substantively related to the focal predictors and outcome.

Call. First, the particular issue customers are calling about could impact agent language and customer satisfaction, so we include dummies to control for the four call categories captured by the firm (*Order*, *Shipping*, *Return*, and *Product*).

Second, the complexity of the call could shape agent language, and their ability to satisfy the customer, so we control for that as well. We take the average of two judges who listened to each call and indicated perceived difficulty or severity of the call on a five-point scale ($r = .72$; *Severity*). In addition, given that complex issues may require more discussion, we control for call length using the total number of words spoken (*Length*).

Third, whether the agent was able to resolve the customer's issue during the call likely impacts how the agent and customer speak, as well as customer satisfaction. To account for this, two judges read each call transcript and indicated whether the customer's main issue had been resolved (1, 0; *Resolved*). Judge disagreements were settled via discussion.

Fourth, rather than the dynamic timing of agent warm and competent language (i.e., *when* language matters), it could be just the overall conversation-level presence of such language that drives any results (i.e., *what* language matters). To account for this, we include controls for agent affective and cognitive language at conversation level.

Agent. An employee's experience could shape how they speak and conversation outcomes, so we control for agent characteristics in two ways. First, to capture organizational experience, we include how many days agents have been with the firm (*Agent Tenure*). Second, to account for direct customer experience, we consider the number of calls they have handled (*Agent Calls*), which is only moderately correlated with tenure ($r = .38$, $p < .05$). These measures help capture unobservable aspects of agent quality or performance (Ng and Feldman 2010). The firm also provided agent gender, which we include as a dummy variable (*Agent Female*).

Customer. Customer attributes can impact satisfaction and purchase, so we control for the two demographics variables provided by the firm, using dummies for which of five geographic regions a customer resides in (*Customer Region*), and for customer gender (*Customer Female*).

Experience with a firm can affect customer satisfaction and behavior, so we control for this in two ways. First, we use the number of days since the customer's first purchase with the firm (*Customer Tenure*). Second, we include their lifetime expenditure with the firm in dollars (*Customer LTV*). Customer attitudes about other aspects of the firm could impact how they interact with the agent, and their satisfaction. To control for this possibility, we also include measures of attitudes towards the website (*Attitude Web*) and shopping experience (*Attitude Shop*), which were captured after the customer satisfaction measure at the end of the call.[4]

Modeling Approach

Functional Data Analysis. To characterize the relationship between the focal dynamic conversational features (e.g., affective and cognitive language) and static conversational outcome (i.e., customer satisfaction), we use semiparametric tools from functional data analysis (FDA; Ramsay and Silverman 1997). Functional data has seen growing applications in marketing to help address dynamic modeling challenges such as predicting motion picture demand (Foutz and Jank 2010), relating moment-to-moment consumer attitudes to TV show judgements (Hui, Meyvis, and Assael 2014), or exploring temporal variations in online chatter and new product performance (Xiong and Bharadwaj 2014).

[4] See Web Appendix Tables A2-A4 for summary statistics and variance inflation factors (VIFs) for the focal predictors and controls. All VIFs fall under the conservative cut-off of 5.

We extend FDA to conversations. We consider time-varying measurement of a conversation feature (e.g., affective or cognitive language) within the n-th conversation as a trajectory $X_n(t)$, $n = 1,...,N$, that is randomly drawn from an underlying stochastic function. The following functional regression relates the static outcome of the interaction y_n to the dynamic language measurement $X_n(t)$,

$$y_n = \alpha + \int_0^1 \beta(t)[X_n(t) - \mu(t)]dt + e_n \tag{1}$$

where α is the intercept, $\mu(t) = \mathbb{E}[X_n(t)]$ the mean function of $X_n(t)$, e_n the i.i.d. Gaussian error term, and $\beta(t)$ the sensitivity curve of interest that characterizes the dynamic impact of a linguistic feature at different moments during a conversation. To meet the requirement that the units of functional analysis have the same duration, we standardize the varied conversation lengths to a common interval [0,1] (Ramsey and Silverman 1997). Therefore, any conclusions should be viewed against the relative progress of a conversation rather than absolute time passed. To account for the potential impact on model estimates due to standardization, we include conversational length in seconds and word count as controls in the main model.[5]

There are also some challenges specific to conversational data (i.e., irregularity and sparsity) that need to be addressed. While virtual stock markets (Foutz and Jank 2010) and continuous user dials (Hui et al. 2014) provide evenly spaced and dense measurements, conversational language occurs over a series of spontaneous conversational turns and tend to be irregularly spaced across time. Further, not every conversational feature (e.g., cognitive words) appears in every turn, resulting in sparse measurement. Except for a handful of calls that contain

[5] Alternatively, one could standardize by conversational turn rather than by time. Compared with the average call length of 371.40 (SD = 238.22) seconds, the mean inter-turn interval of 0.26 (SD = 0.53) seconds is negligible and so standardization by time is preferred.

close to 100 measures of some language features, most interactions have 10 to 30 turn-level measurements. Consequently, functional regression for conversation must be able to handle the irregular and sparse presence of language features (see Web Appendix Figures A3 and A4).

Our dynamic modeling approach addresses these challenges. We consider a dynamic unstructured language feature as a continuous trajectory $Z_n(t)$ over the course of conversation n. Across multiple conversations, we obtain a sample of measured trajectories assumed to be independently drawn from an underlying stochastic function, with unknown mean function $\mu(t) = \mathbb{E}[Z_n(t)]$ and variance function $\Sigma(t_1,t_2) = \mathrm{Cov}[Z_n(t_1),Z_n(t_2)]$. Due to measurement errors arising from using language dictionaries, the actual observation for the m-th measurement, $m = 1,\ldots,M_n$, of the n-th conversation is given by

$$X_n(t_m) = Z_n(t_m) + \varepsilon_n(t_m) \tag{2}$$

where t_m indicates the time of the sequential conversational turn at which the measurement was taken, and the measurement error ε_n is i.i.d. drawn from $N(0,\sigma^2)$. In call n, the M_n measurements are irregularly-spaced and sparse. We assume M_n is exogenous and control for its effect in our model.

For the focal functional predictors (agent affective and cognitive language), we apply scatterplot and surface smoothing, both via local linear regression, to estimate mean and covariance functions respectively (Yao, Muller, and Wang 2005; Wang, Chiou, and Muller 2016; Chen et al. 2016).[6] We use the entire sample simultaneously in the smoothing procedure to allow information shrinkage across observations to accommodate the sparseness discussed above.

After smoothing, we apply Karhunen-Loève expansion to obtain eigen components of the

[6] For both the smoothed mean and covariance functions, we apply the commonly-used Gaussian kernel and obtain the smoothing bandwidth via the generalized cross-validation bandwidth selection (Speckman 1988).

conversations, $\{X_n(t)\}_{n=1}^{N}$, namely,

$$\Sigma(t_1,t_2) = \sum_{i=1}^{\infty} \lambda_i \phi_i(t_1)\phi_i(t_2) \tag{3}$$

and so

$$X_n(t) = \mu(t) + \sum_{i=1}^{\infty} \omega_{ni}\phi_i(t) + \varepsilon_n(t) \tag{4}$$

where $\phi_i(t)$ is the i-th eigen function, λ_i the associated eigen value, and ω_{ni} the i-th eigen score of the n-th conversation. If we expand the unknown $\beta(t)$ curve onto the same eigen bases,[7]

$$\beta(t) = \sum_{i=1}^{\infty} b_i \phi_i(t) \tag{5}$$

thanks to orthogonality, the functional regression in (1) can now be simplified to

$$y_n = \alpha + \sum_{i=1}^{\infty} b_i \omega_{ni} \approx \alpha + \sum_{i=1}^{I} b_i \omega_{ni} \tag{6}$$

In the above, the truncation I, or the actual number of eigen components to appear in the regression, is determined using AIC. We also tested metrics such as BIC and leave-one-out cross-validation, and saw almost identical truncations across language features.

The above approach allows us to examine the relationship between the dynamic moments (turns) of our focal dynamic predictors (agent affective and cognitive language) and the static outcome (customer satisfaction). When there are multiple functional predictors and scalar controls, we can describe a generalized functional regression as follows,

$$E\left[y_n \,\middle|\, \{X_{ln}\}_{l=1}^{L}, \{W_{jn}\}_{j=1}^{J}\right] = g^{-1}\left(\alpha_a + \sum_{l=1}^{L} \int_0^1 \beta_l(t)[X_{ln}(t) - \mu_l(t)]dt + \sum_{j=1}^{J} \gamma_j W \right. \tag{7}$$

[7] Alternatively one could use Riemann sum to remove the integral without assuming identical bases for $\beta(t)$. But doing so would introduce numerical errors into the estimation and burden the subsequent model regularization with many additional variables.

where L and J denote the number of functional predictors and scalar controls respectively, W_{jn} is the j-th scalar control for the n-th call, γ_j represents the regression coefficients, and $g(\cdot)$ indicates the link function for a nonlinear dependent variable. Besides using agent observables as controls, we capture unobserved agent heterogeneity with a random intercept α_a for every agent.

Applying the smoothing procedure and Karhunen-Loève expansion to the data, we obtain a simplified generalized regression as follows,

$$E\left[y_n \,\middle|\, \{X_{ln}\}_{l=1}^{L}, \{W_{jn}\}_{j=1}^{J}\right] = g^{-1}\left(\alpha_a + \sum_{l=1}^{L}\sum_{i=1}^{I_l} b_{li}\omega_{lni} + \sum_{j=1}^{J}\gamma_j W_{jn}\right). \tag{8}$$

where I_l for function variable $X_l(t)$ is determined by the truncation criterion discussed above.

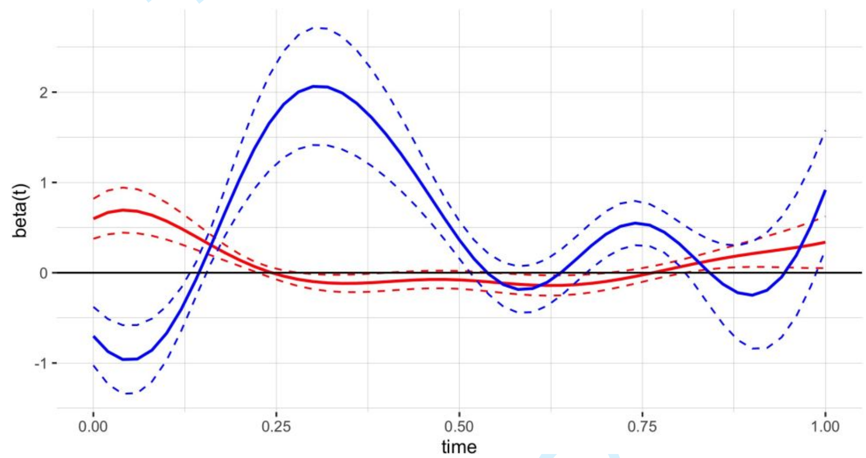
Main Results

Figure 2 presents the key results. Functional regression results are depicted as a beta curve ($\beta_l(t)$ or "beta(t)"), plotting the moment-to-moment beta coefficients for the focal affective and cognitive language predictors over conversational time. Model 1 shows the relationship between affective and cognitive language and customer satisfaction, and Model 2 presents the same results after adding the controls. When the pointwise 95% confidence interval (dotted line) is above (below) zero for one of these language features, that feature has a positive (negative) relationship with the customer satisfaction outcome at that particular point in conversational time, allowing one to interpret *when* affective and cognitive language matter. For example, model results reveal that approximately 12.5% into a service conversation, affective language (red line) has a positive and significant beta coefficient of 0.5, and cognitive language (blue line) has a negative and significant beta coefficient of 0.3. The relative scale of the coefficients signals their relative importance across both predictors and moments.

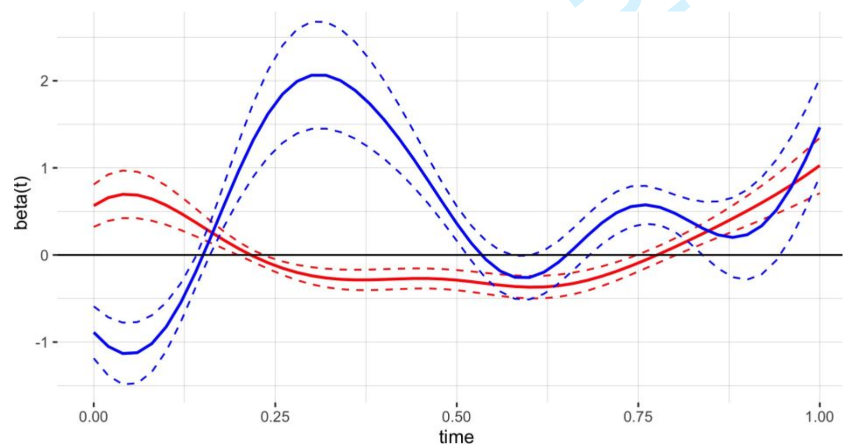
As predicted, customers are more satisfied when agents use more affective language at the beginning and end of conversations. But affective language is not beneficial during the middle of the call.

Figure 2: Agent Language and Customer Satisfaction

Model 1 (no controls)



Model 2 (Model 1 + controls)



Red lines: Affective Language; Blue lines: Cognitive Language
Dotted lines: pointwise 95% confidence intervals

Cognitive language results are quite different. Speaking more rationally at the beginning of conversations appears to be costly, but customers are more satisfied when agents use more cognitive language in the middle of the conversation.

Taken together, these findings suggest that affective and cognitive language are *both* linked to positive satisfaction outcomes, but at *different times* during an interaction.[8] Customers were more satisfied when agents use warm language at the start and end, but cognitive language primarily in the middle. Further, a comparison of the optimal dynamics of agent language (Figure 2) to actual language use (Figure 1) shows that agents are not using language this way currently, casting doubt on the notion that these patterns are somehow already known and in use.

Additional Unstructured Controls

While the 22 factors controlled for are more than prior conversation dynamics research in marketing (e.g., Singh et al. 2018; Singh et al. 2020), one can always wonder about additional possible sources of endogeneity. We test causality through four experiments, but to further explore the field data, we also consider unstructured text and voice controls.

One of the benefits of unstructured data is the ability to control for a wide range of features. Aspects of language, vocal features (e.g., pitch), and, in other data, images, that vary across conversational moments (e.g., turns) can now be measured. As such, one can consider myriad factors that might help explain a focal relationship, and by including them in the model, test potential alternative explanations (Berger, van Osselaer, and Janiszewski 2024).

[8] Corroborating prior research (e.g., Marinova et al. 2018;), the size of cognitive language's positive coefficient supports the importance of a competence-oriented approach. That said, the present study reveals *when* in conversation conveying competence is important (e.g., middle), and that its use can be determinental if used at the wrong conversational moments (e.g., start).

That said, this benefit comes with a downside. There are hundreds, if not thousands of potential unstructured data dimensions researchers could include, and as more variables are considered, *overfitting* becomes a problem. Further, it is problematic to include controls due only to their availability (Clarke 2006; Spector and Brannick 2011).

Nonetheless, to further control for possible sources of endogeneity, we apply a machine-learning method, Group-Lasso (Yuan and Lin 2006; Meier et al. 2008; Yang and Zou 2015), that attempts to incorporate as many of the unstructured controls as appropriate while preventing overfitting. The Group-Lasso regularization helps avoid the path-dependency problem in conventional stepwise regression (e.g., Foutz and Jank 2010), and allows for *group-wise* variable selection as the selection of functional variables corresponds to selecting from the L groups of eigen scores in (8) (see Web Appendix B for more details).

For this wide data exercise, we consider an additional 28 text and voice controls (see below), which equal up to 111 potential additional control parameters after calculating their eigen components to account for moment-to-moment dynamics.

Dynamics of Other Major Agent Language Features. First, beyond affective and cognitive language, other moment-to-moment features of employee language may shape how customers perceive or speak to them. To attempt to control for this, we include dynamic, turn-level measures of LIWC's other main psychological process dictionaries (e.g., *Social processes, Perceptual processes, Drives, Temporal perspective, and Informality*; Pennebaker et al. 2015).

Dynamics of Agent Paralanguage. In addition to what was said, one could wonder whether how things were said (i.e., paralanguage) might drive the effects. We attempt to control for dynamic acoustic features linked to persuasion (Van Zant and Berger 2020) at the turn level using phonetics software (*Pitch* and *Intensity*; Boersma and van Heuven 2001) applied to the original audio call recordings.

Dynamics of Customer Affective and Cognitive Language. Agents might mimic or repeat recent customer language, which could shape agents' affective and cognitive language (the focal IVs). To account for this possibility, we attempt to include the customer's own affective and cognitive language over the course of the conversation as dynamic controls.

Dynamics of Other Major Customer Language Features. Beyond affective and cognitive language, other moment-to-moment aspects of customer language may shape how employees speak, so we attempt to control for these using turn level measurement of the same psychological process dictionaries used for employee language (i.e., *Social processes, Perceptual processes, Drives, Temporal perspective, and Informality*).

LDA Topics. To account for a more fine-grained mixture of topics than the five call categories provided by the firm, we use customer language to uncover the hidden mixture of topics via topic modeling (i.e., latent Dirichlet allocation (LDA); Blei, Ng, and Jordan 2003). Standard pre-processing included stemming related words (e.g., walk, walked, or walking = walk) and removing punctuation and numbers. Results were robust to the inclusion or exclusion of infrequent words and stop words. We followed suggested practices and prior research (Blei 2012; Chang et al. 2009) in determining the number of topics. We examined 5-15 topic solutions, and perplexity fit measures revealed a peak (lower perplexity) at 13 topics, so we attempted to include the 13 topic model results as additional controls.

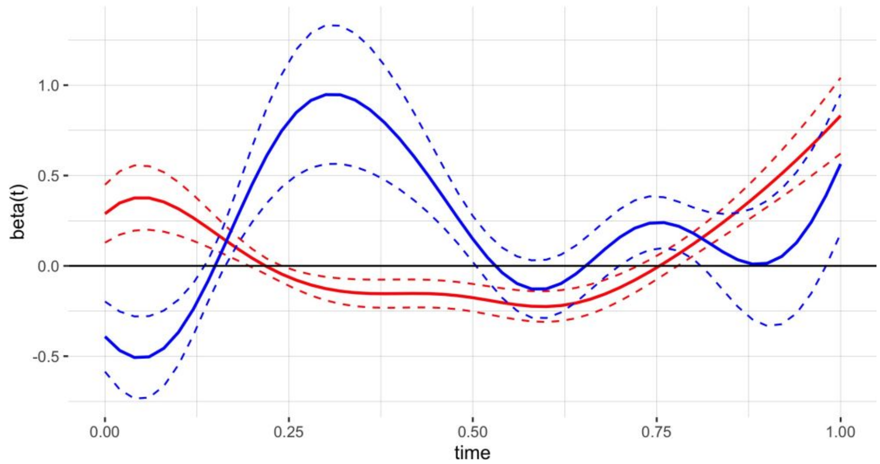
Moment-to-Moment Linguistic Synchronicity. To further isolate the dynamic impact of agent language, we further consider how it may be shaped by customer language over the conversation. How someone speaks can impact their conversation partner, but also can reflect what the conversation partner said previously (Goffman 1981; Grice 1991; Zhang et al. 2020). To control for these aspects, we use a moment-to-moment measure of linguistic synchronicity (*Synchronicity)*. Specifically, following Zhang, Wang, and Chen (2020) we create a synchronicity measure using the R^2 of the moment-to-moment regression from customer language on agent language. See Web Appendix Figure A2 for details.

Model. As discussed, while these additional unstructured text and voice controls help further assess robustness to omitted control endogeneity, given the large number of unstructured controls and their moments (N = up to 111 additional control parameters), one could worry about overfitting. Consequently, we use Group-Lasso machine learning to penalize out unstructured controls that impede model fit and inference (see Web Appendix B for method details). The method selected 23 additional unstructured control parameters in this extended model (Model 3), in addition to the 22 controls considered in Model 2.

Results. Results of Model 3 (Figure 3) are highly similar to the functional forms observed in Models 1 and 2. Specifically, affective language is beneficial at the start (25%) and end (25%), but not in the middle (50%) of these conversations. In contrast, cognitive language is costly at the start, beneficial in the middle, and null for most of the conversation's end.[9]

Figure 3: Study 1 Model 3 (Model 2 + unstructured controls after Group-Lasso)

[9] Table A7 in the Web Appendix presents parameter estimates for the focal predictors, structured controls, and additional wide data unstructured controls across all three Study 1 models.



Red lines: Affective Language; Blue lines: Cognitive Language
Dotted lines: pointwise 95% confidence intervals

Discussion

Overall, results suggest that the relationship between agent language and customer satisfaction depends on when in the conversation it occurs. Consistent with our theorizing, rather than a more cognitive, competence related language style being beneficial throughout, it is mainly helpful in the middle of conversations. Warmer, more affective language is beneficial at the conversation's start and end. Results are robust to the inclusion of over 40 traditional and unstructured (text and voice) control variables. While it is difficult to rule out omitted variable endogeneity in conversational data (Reece et al. 2022; Zhang et al. 2020), considering a wide variety of factors potentially linked to our focal IVs and customer satisfaction helps mitigate such concerns.

Robustness. We also performed several additional robustness tests (see Web Appendix B for detailed results). First, we tested robustness to a different outcome variable: purchases. Results follow similar functional forms (e.g., affective language beneficial at the start, cognitive language in the middle), suggesting the benefit of our dynamic approach may extend to important downstream behaviors.

Second, results are robust to using other relevant language dictionaries from prior research (e.g., “relating” vs. “resolving” from Marinova et al. 2018; Singh et al. 2018).

Third, the link between affective language and customer satisfaction is robust to considering only positive or negative language, but is more strongly driven by positive language.

Relative Contribution of Affective and Cognitive Language. While results thus far suggest conversational moments when affective and cognitive language are each beneficial, one might wonder which language is more important, “overall.” To consider this question, we compare the proportions of positive versus negative areas under the beta curve for each functional feature. Results indicate that the majority of both affective (65.37%) and cognitive language (80.62%) contributions are positive, if emphasized at the right time. The larger positive contribution area for cognitive language suggests that, if the timing of these two speaking styles is optimized, cognitive language will make a greater contribution. The larger negative area for affective (34.63%) than cognitive (19.38%) language suggests it is particularly important for agents to know when to speak to customers more affectively (i.e., start and end).

Benchmarks and Simulations. We also investigated whether our approach performs better than competing benchmarks (see Web Appendix B). Our dynamic model yields stronger in-sample and out-of-sample predictions than (1) traditional “what” analysis that does not account for dynamics at all, (2) a “what” analysis that includes the “sensing, seeking, and settling” conversational stages offered in Marinova, Singh, and Singh (2018), (3) our functional model including all additional unstructured text and voice controls without consideration of model overfitting, and (4) a model ignoring the agent heterogeneous effect. Taken together, this suggests our approach offers superior predictive performance relative to previous models.

To further test these ideas, we performed a series of simulations comparing our model with various alternatives in what language is used when. Results underscore the benefits of using *both* affective and cognitive language, rather than only one, and of considering *when* to use each of these approaches over the course of a conversation beyond merely *what* language is used overall. See Web Appendix B for detail.

STUDY 2: AIRLINE FIELD DATA

While the initial results are intriguing, one might wonder whether they are driven by the specific firm, industry, or customer satisfaction measure used. To test generalizability, we worked with a major U.S. airline to acquire an additional randomly selected (by the firm) dataset of 204 customer service calls (11,548 conversational turns). The airline captured willingness to recommend at the end of the call, a measure widely used to assess customer satisfaction (e.g., Keiningham et al. 2007; van Doorn, Leeflang, and Tijs 2013).

Model 1 examines this outcome as a function of agent affective and cognitive language dynamics, and Model 2 used a similar set of structured controls as in Study 1. As in Study 1, we created a control for *Call Complexity* (length in words). The airline was not able to provide customer or agent observables, but provided their measure of *Call Category* (which of four *Departments* the calls were routed to), and whether customers received an *Exchange* or *Refund*.[10] Model 3 includes additional unstructured controls that further add to model fit and inference.

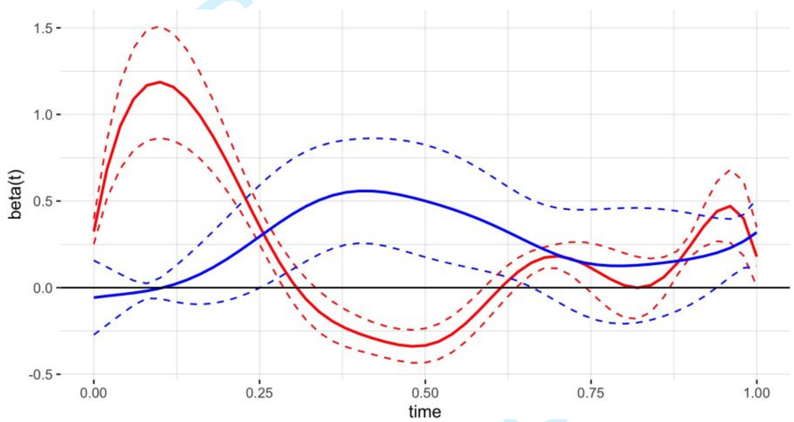
Results. Even exploring a different company, in a different industry, results are similar (Figure 4). Customers were more willing to recommend the airline when agents used more

[10] The firm blinded the researchers to the Category and Department names. They are represented only as numbers.

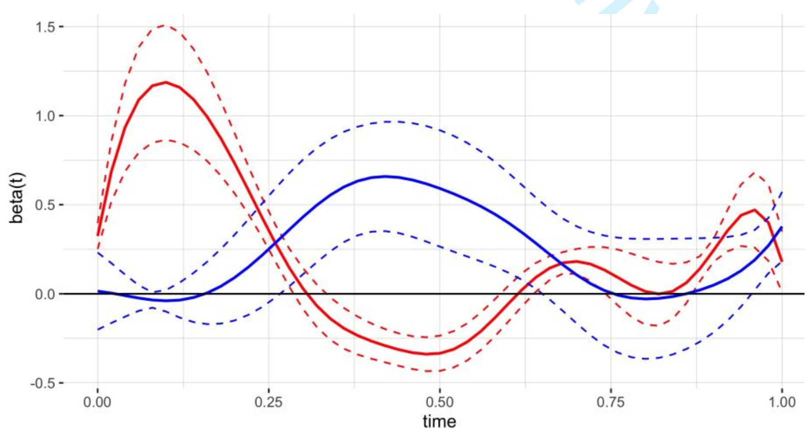
affective language at the start and end of the conversation, but more cognitive language in the middle. Further, as shown in the retailer data, airline agents do not already follow the estimated sensitivity curves (Figure 4 vs. Web Appendix Figure C1), casting additional doubt on the notion that these patterns are somehow already known and practiced. Regression coefficients for predictors and controls for all three models are presented in Web Appendix Table C1.[11]

Figure 4: Study 2 Agent Language and Willingness to Recommend

Model 1 (no controls)

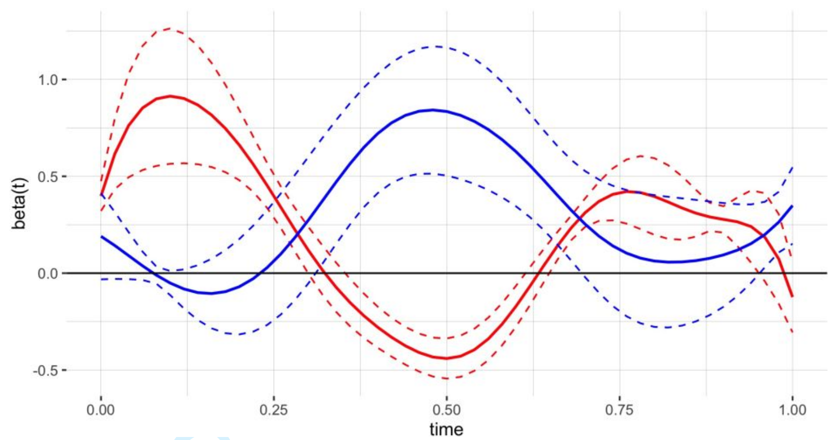


Model 2 (Model 1 + controls)



[11] We also present the results of an analysis that attempts to pool the Study 1 and Study 2 data in the Web Appendix.

Model 3 (Model 2 + unstructured controls after Group-Lasso)



Red lines: Affective Language; Blue lines: Cognitive Language
Dotted lines: pointwise 95% confidence intervals

STUDY 3: INITIAL CAUSAL TEST ACROSS NATURALISTIC STIMULI

Finding the same results across two different field datasets underscores their validity and generalizability. That said, one could wonder whether the effects are causal. Including a large number of control variables helps cast doubt on many alternative explanations, but it's still possible some unobserved factor could explain the results. Alternatively, perhaps agents infer the customer's satisfaction early on in the conversation, and this shapes their subsequent language (i.e., reverse causality).

To more directly test when language matters, Study 3 manipulates it. We vary agent language to test whether, compared to the strategy recommended in prior research (i.e., emphasizing competence throughout; Kirmani et al. 2017; Li et al. 2019; Marinova et al. 2018), the dynamic strategy recommended by our conceptualization (and supported by Studies 1 and 2, i.e., using more affective language at the beginning and end) boosts customer satisfaction. The experimental approach used in this and subsequent studies also helps assess the validity of the functional regression modeling approach using a more familiar method.

To maximize external validity, we use five different conversations from the Study 1 field data to assess robustness to stimulus sampling. This study was preregistered (https://aspredicted.org/M1K_4VC). All experiments used the same exclusion criteria, and replicate without the exclusion (see Web Appendix D). Achieved power after exclusion was greater than 85% (α = 5%) for all experiments.

Method

Participants (N = 686, Prolific) were randomly presented with the full transcript of a version of one of five real service conversations sampled from Study 1. To approximate the topic distribution in the field data, we sampled across all of the firm's call topics, and included calls related to returns, orders, shipping, and product (see Web Appendix Table A3).

The only difference between conditions was agent language. In the control condition, participants saw the original conversation transcript, edited to remove personally identifiable information (e.g., customer's address and company name). In the dynamic treatment condition, employee language was adjusted based on the dynamic findings of Study 1 and 2. Specifically, agents used warmer, more affective language (e.g., words and phrases like "feel," "sorry," and "no worries," all adapted from the LIWC affective dictionary) in the first and last 25% of each conversation. See Web Appendix D for full stimuli and affective language LIWC scores by condition.

After reading one of the ten conditions (2 (language: control vs. treatment) x 5 (conversational variant: return 1, return 2, order, shipping, product)), participants were asked "How satisfied would you be with the employee?" (1 = not at all, 7 = very much).

Results

As predicted, across a range of real customer service conversations, using our dynamic language recommendation boosts customer satisfaction ($M_{treatment}$ = 5.10, SD = 1.81 vs. $M_{control}$ = 4.61, SD = 1.86; $F(1, 684) = 12.45$, $p < .001$, $\eta^2_p = .02$).

Results remain the same controlling for conversation variant and its interaction with language condition ($F(1, 676) = 17.21$, $p < .001$, $\eta^2_p = .03$). Further, the benefit of adding more affective language to the start and end did not vary across the five conversations (interaction $F(4, 676) = .62$, $p = .645$). See Web Appendix D for condition means for all five stimuli.

Discussion

An externally-valid experiment, sampling a variety of real customer service interactions, provides direct causal support for our theorizing. Consistent with our suggestion, and with Studies 1 and 2, using more affective language at the start and end boosted customer satisfaction.

Ancillary analyses also cast doubt on the notion that the effects could be driven by *what* rather than *when*. If the condition that used more affective language at the start and end also used more affective language overall, maybe it is the greater amount of affective language used, rather than when it occurred, that is increasing customer satisfaction. To test whether this alternative can explain the results of Study 3, we control for the propotion of affective (and cognitive) language in each stimuli variant as covariates. The effect for our dynamic treatment remains significant ($F(1, 682) = 124.04$, $p < .001$, $\eta^2_p = .15$).[12]

STUDY 4A: CONTROLLED STIMULI

While Study 3 provides direct causal evidence using a range of real conversations, the idiosyncratic and complex nature of natural conversation makes it difficult to maintain strong experimental control (Reece et al. 2022). Consequently, Study 4 provides a simpler, more controlled language manipulation.

Method

Participants (N = 146, MTurk) were randomly assigned to one of two versions of a simple scenario based on the field data conversations. Shipping related issues were common in Study 1 (49 conversations) and were perceived to be approximately average in severity ($M_{shipping} = 2.84$, $SD = .91$ vs. $M_{all} = 2.61$, $SD = .94$), so participants imagined calling an online retailer, and read a conversation in which they asked the customer service agent for shipping help.

The only difference between conditions was the agent's language. As recommended by prior research, in the all-cognitive condition, the agent used cognitive language throughout (i.e., a "competent-competent-competent" sequence). In the dynamic condition, agent language followed the findings of Study 1 and 2. Specifically, in the first and last 25% of the conversation, cognitive language was replaced with more affective language from the LIWC affective

[12] Note that our modeling results (Studies 1 and 2) already account for, and our simulations (Web Appendix B) directly test, the effects of overall agent use of affective language, and thus cast doubt on this alternative. We also carefully control for the total amount of warm, affective language used in Study 5.

dictionary (i.e., a "warm-competent-warm" sequence). In the all-cognitive condition, for example, the agent started by saying "Hello. How might I assist you today?", while in the dynamic condition they used the warmer "Hello. I hope you're enjoying this fine day?"[13] See Web Appendix D for full stimuli.

Then, participants completed the key dependent variable (i.e., customer satisfaction, "How satisfied are you with the agent?"; 1 = not at all, 7 = very much). To replicate the Study 1 retailer's satisfaction measure, we also asked "How helpful was the agent?" (1 = not at all, 7 = very much).

Results

As predicted, changing agent language based on our dynamic recommendation (i.e., more affective language at the start and end) improved customer satisfaction ($M_{dynamic} = 6.30$, $SD_{dynamic} = .73$ vs. $M_{all\ cognitive} = 5.87$, $SD_{all\ cognitive} = .89$; $F(1, 144) = 10.25$, $p = .002$, $\eta^2_p = .07$). It also led agents to be perceived as more helpful ($M_{dynamic} = 6.14$, $SD_{dynamic} = .88$ vs. $M_{all\ cognitive} = 5.84$, $SD_{all\ cognitive} = .93$; $F(1, 142) = 4.07$, $p = .046$, $\eta^2_p = .03$).

Discussion

[13] While one might wonder whether the dynamic language condition recommended by our model seemed less typical, expected, or standard, this was not the case. There was no difference in perceived language typicality across conditions ($F < 1$ using the three-item measure from Kronrod, Grinstein, and Wathieu 2011), casting doubt on this alternative.

Controlled manipulation of the language used at different conversational stages provides further causal support. Consistent with our theorizing, and with the results of the first three studies, dynamic "warm-competent-warm" language boosted customer satisfaction over previously recommended approaches prioritizing competence throughout (i.e., "competent-competent-competent").

STUDY 4B: COMPARISON TO OTHER LANGUAGE SEQUENCES

While the results of Study 4A are supportive, one could wonder whether other sequences of affective and cognitive language might be more beneficial. To test this possibility, Study 4B extends Study 4A, adding six additional carefully controlled language sequence conditions. This study was preregistered (https://aspredicted.org/Y2Y_SZC)

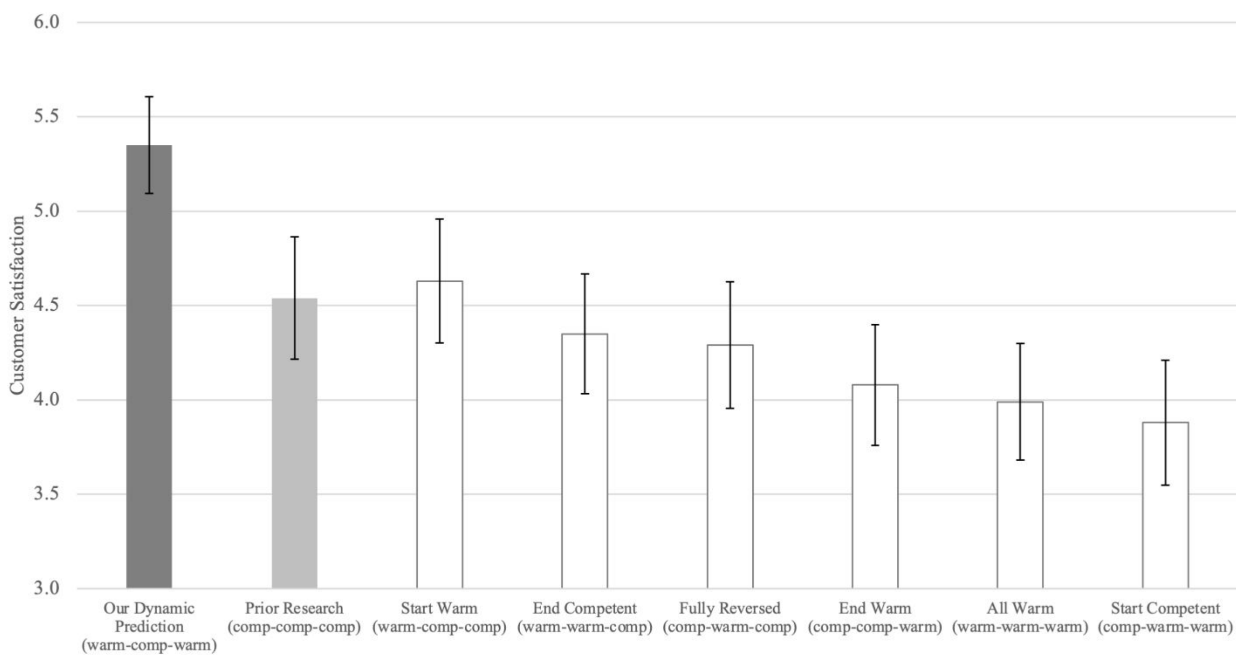
Participants (N = 603, Amazon Mechanical Turk) were randomly assigned to one of eight versions of the base stimuli from Study 4A. The first two conditions were identical to Study 4A. The third and fourth conditions take our recommended "warm-competent-warm" approach and shift either the first or last period to be competent instead (i.e., "competent-competent-warm" or "warm-competent-competent"). The fifth condition tries warmth throughout (i.e., "warm-warm-warm") and the sixth condition fully reverses our suggestion (i.e., "competent-warm-competent"). Both of these conditions use warmer, more affective language (e.g., "I've been frustrated locating it myself") in the middle of the conversation. Notably, the fully reversed condition uses the same total amount of agent warm and competent language, ruling against the possibility that this can drive the effect.[14] The seventh and eighth conditions include the final two permutations for completeness (i.e., "competent-warm-warm" and "warm-warm-competent"; see Web Appendix D for stimuli).

Results indicate that language based on the dynamic model's recommendation improved customer satisfaction (M = 5.35, SD = 1.60) relative to all other conditions (Figure 5). This includes the competence throughout recommendation of prior research (M = 4.54, SD = 2.02; $F(1, 146) = 7.20$, $p = .008$, $\eta^2_p = .05$) as well as warm only at the start (M = 4.63, SD = 2.06; $F(1, 143) = 5.47$, $p = .021$, $\eta^2_p = .04$), warm only at the end (M = 4.08, SD = 2.01; $F(1, 146) = 17.95$, $p < .001$, $\eta^2_p = .11$), warm throughout (M = 3.99, SD = 1.93; $F(1, 147) = 21.75$, $p < .001$, $\eta^2_p = .13$), competence-warmth-competence (M = 4.29, SD = 2.10; $F(1, 145) = 11.65$, $p < .001$, $\eta^2_p = .07$), competence-competence-warmth (M = 4.08, SD = 2.01; $F(1, 146) = 17.95$, $p < .001$, $\eta^2_p = .11$), and warmth-competence-competence (M = 4.63, SD = 2.06; $F(1, 143) = 5.47$, $p = .021$, $\eta^2_p = .04$).[15] These findings underscore the notion that the specific dynamic sequence from our theorizing is superior to a variety of alternative sequences, and further supports our prediction that *when* language is used matters (rather than merely *what* language is used).

Figure 5: Comparison Against Various Alternatives (Study 4B)

[14] The proportion of overall agent words in the fully reversed "competent-warm-competent" condition are the same as in our dynamic treatment condition ("warm-competent-warm") for both affective (8.9% vs. 10.6%; $\chi^2 = .005$, $p = .778$) and cognitive language (22.2% vs. 21.3%; $\chi^2_{cognitive} = .040$, $p = .841$).

[15] As in Study 4, results also replicate using the Study 1 retailer's satisfaction measure "How helpful was the agent?". Our dynamic treatment condition again outperformed the recommendation of prior research ($M_{dynamic} = 5.54$, $SD_{dynamic} = 1.58$ vs. $M_{all\ cognitive} = 4.87$, $SD_{all\ cognitive} = 2.02$; $F(1, 146) = 5.07$, $p = .026$, $\eta^2_p = .03$) and all six other conditions (all *p*s < .02; all $\eta^2_p > .03$).



Note: Error bars represent 95% confidence intervals. Text between parentheses describes the manipulated sequence of more affective (warm) or more cognitive (comp) agent language for each condition.

STUDY 5: REPLICATION AND ROBUSTNESS

Studies 1, 2, 3, 4A and 4B offer evidence that, beyond *what* language agents use overall (i.e., conversation-level use of warm language), *when* agents use it matters (i.e., at the start and end). Study 4B, for example, offers a particularly conservative test through the fully reversed "competent-warm-competent" condition that uses the same overall amount of warm language as our dynamic treatment ("warm-competent-warm"), but at the wrong time. Study 5 extends this approach further, testing our dynamic treatment using a "competent-warm-competent" control that uses exactly the same number and proportion of warm words across these two conditions.

Method

We randomly assigned participants (N = 154, Prolific) to one of two versions of a simple airline service scenario based on the Study 2 field data conversations. This study was preregistered (https://aspredicted.org/YL7_9LY).

The only difference between conditions was the agent's language. For our dynamic treatment condition, agent language once again followed our recommended "warmth-competence-warmth" sequence. In the fully reversed control condition, the agent used more cognitive language at the start and end, and more affective language in the middle (i.e., a "competence-warmth-competence" sequence). In the control, for example, the agent used warmer language in the middle "I'm just hoping to share something that might be alright for you", while in the dynamic condition they used more cognitive, competent language at this time "I'm just trying to find something that might work for you". To fully control for the overall count and proportion of affective and cognitive language that agents used, we made sure they were identical across the conditions. See Web Appendix D for full stimuli.

Participants completed the same customer satisfaction dependent variable as in all prior experiments.

Results

As predicted, even though it used the exact same number and proportion of warm and competent agent words overall, agent language based on our dynamic recommendation (i.e., warmth-competence-warmth) improved customer satisfaction ($M_{dynamic} = 5.74$, $SD_{dynamic} = 1.26$ vs. $M_{fully\ reversed} = 5.06$, $SD_{fully\ reversed} = 1.38$; $F(1, 152) = 9.98$, $p = .002$, $\eta^2_p = .06$).

GENERAL DISCUSSION

Language impacts a range of consumer interactions. But while a great deal of research has examined customer service language and other marketing dialogues (e.g., social media conversations; Berger and Schwartz 2011; Ordenes and Grewal 2017), *when* different language features matter in conversation has received less attention.

To address this gap, we offer an approach that examines how language at different moments of an interaction relates to important outcomes. As an initial demonstration, we applied it to the two most important dimensions of person perception: warmth and competence. While existing research suggests that either competence (in customer service) or warmth (in everyday interpersonal relations) should take primacy, our approach suggests a more dynamic perspective may be beneficial. Consistent with this, six studies find that "bookending" the efficient, competent addressing of customer needs with warmer, more affective rapport building at the start and end of service interactions increases customer satisfaction. Finding the same results in the lab and two field settings, across a range of naturalistic and controlled stimuli, using different topical contexts and words, and different dependent measures (i.e., customer satisfaction, helpfulness, purchase behavior, word of mouth intentions) speaks to their generalizability. Simulations (see Web Appendix B) speak to the ceiling of the potential impact of these effects.

Importantly, these results go beyond existing research and practice. Launching straight into the competence-oriented language endorsed by prior research may hurt customer satisfaction and purchase, as may using only a warmth-oriented approach. Instead, results suggest that agents should use warmer language at the start and end of conversations than they do currently, and generally avoid more cognitive, competence-oriented approaches during these periods. Language like "My pleasure. Take care now," should be used at the end of conversations, for example, rather than language such as "I'm glad we could solve that for you. Bye now."

Our modeling approach also helps address three major challenges in examining moment-to-moment dynamics in communications–irregularity, sparsity, and high dimensionality (e.g., wide data unstructured text and voice controls). Language measurement is often irregular and sparse, so we modeled the time-varying data as random trajectories realized from smooth underlying functions. We used Group-Lasso machine learning to select additional unstructured controls that enhanced, rather than impeded, model fit and inference.

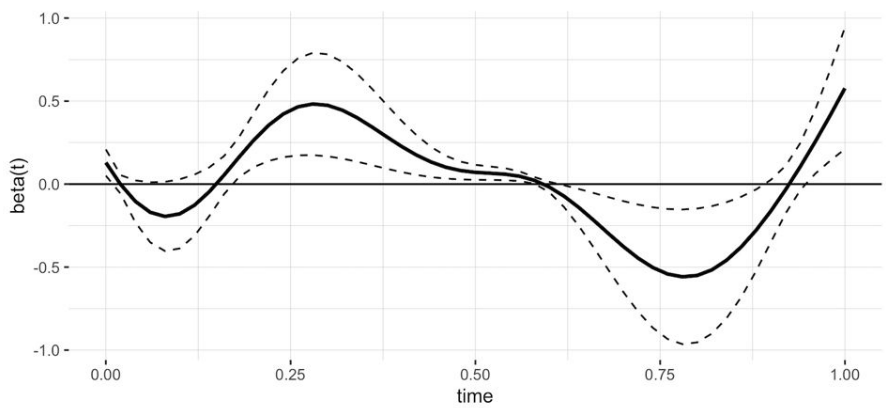
Applications to Other Linguistic Features

We focused on affective and cognitive language, but our method can be applied to any language (or paralanguage) feature. Take questions. Prior research suggests asking questions can be beneficial (Huang et al. 2017) because it signals interest (Drollinger and Comer 1997). Consumers also believe that asking questions is important, making it a common feature of scales used to evaluate employee performance (Drollinger et al. 2006; Ramsey and Sohi 1997).

But while our main dataset (Study 1) replicates prior findings that customers are indeed more satisfied when agents ask more questions overall (b = .13, p = .010), is asking questions good at any point in the conversation? Or might it be more beneficial in certain parts?

To illustrate how our method can test such ideas, we run our functional model with agent question-asking as the focal dynamic predictor of customer satisfaction. Results indicate that the positive relationship between customer satisfaction and question asking depends on *when* agents do so (Figure 6). While asking questions is not helpful in the first 15%, doing so is beneficial when used between 15% and 57% of the interaction, and can even be costly at 60-85% of the way through. This suggests agents might best emphasize questions after the customer has a chance to describe their needs.

Figure 6: Agent Question Asking and Customer Satisfaction

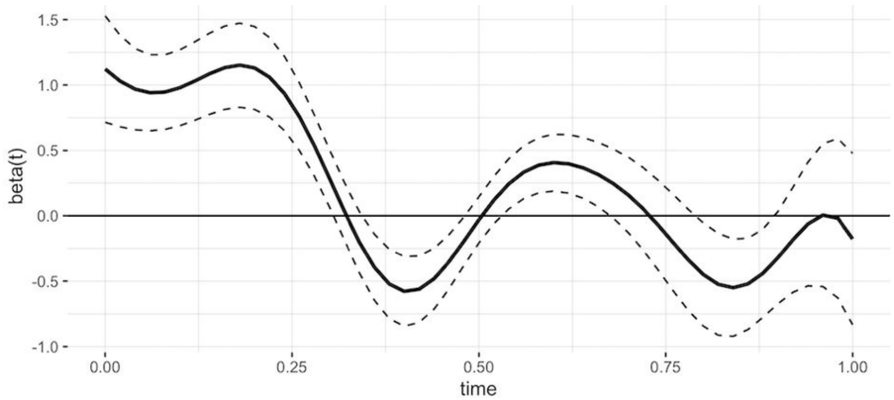


Dotted lines: pointwise 95% confidence intervals

To further explore the method's value, we also looked at pronouns. Research suggests that first person singular ("I") pronouns make agents seem more agentic and empathetic (Packard et al. 2018), and a traditional conversation level analysis of the Study 1 field data replicates the finding that first person singular pronouns are positively related to customer satisfaction overall (b = .051, p = .040). But are these pronouns necessarily important throughout a conversation?

Running the same model with agent first person singular pronouns as the main dynamic predictor finds that their benefit mostly occurs at the beginning of conversations (Figure 7). This is the same period when warm, affective language is beneficial. In contrast, first person singular pronouns may be costly for a brief period when cognitive language matters (i.e., the middle of the conversation). This pattern suggests that first person perspective may be more important when conveying warm empathy ("I'm sorry") than signaling competent agency ("I'll fix it"). Competence might be better achieved by using more objective voice (e.g., third person).

Figure 7: Agent First Person Singular Pronouns and Customer Satisfaction



Dotted lines: pointwise 95% confidence intervals

Overall, these examples further underscore the potential value of examining language dynamics, demonstrating not only whether the words we use matter, but *when*.

Substantive Implications, Limitations, and Future Research

Our findings have clear implications for researchers and managers. For researchers, our approach offers a way to move beyond just *whether* certain language features matter to *when* they matter. This method expands the toolkit available to researchers who use text analysis to understand consumer behavior (Berger et al. 2020; Humphreys and Wang 2018). It could easily be applied to paralanguage (Luangrath, Peck, and Barger 2017) or non-verbal communications, and other long-form language contexts (e.g., advertising copy, movie scripts, or online reviews).

Managers can use the approach to understand not only what language to use, but when to use it (see Table 1 for examples). When trying to design more effective chatbots, for example, understanding when to prioritize different language features and non-verbal cues (e.g., tone, pitch, pauses) should make these conversational technologies more effective.

Table 1: Managerial Training Examples of Service Agent Language

Conversational Moments	Language Style	Example Agent Turns (adapted from Study 1 data)	Link to Outcomes
Opening	**More Affective / Warm**	**Who do I have the pleasure of speaking with today?**	**Positive**
	More Cognitive / Competent	How might I assist you today?	Negative
Middle	More Affective / Warm	I'm sorry, do you mind sharing your address again?	Negative
	More Cognitive / Competent	**And could you verify your address again?**	**Positive**
Closing	**More Affective / Warm**	**Sure. Glad I could help. Call us back and we'll take care of you.**	**Positive**
	More Cognitive / Competent	Of course. Not a problem. Call us back if you need anything else.	Negative

Examples of language style recommended by the present research are in* ***bold**.*

We accounted for agent, customer, and firm level factors, but as with most field data, our estimates remain subject to potential endogeneities due to unobserved factors. The temporal sequence of our language predictors and outcomes makes reverse causality seem unlikely, and four experiments using both naturalistic and controlled stimuli support causality. But future research could use field experiments to further test external validity.

We focused on effects of language over time, but future work could delve more deeply into the mechanisms behind these effects. We theorized, for example, that warmer, more affective language should be beneficial at the start because it helps establish a warm, relational base before competently addressing the customer's needs. Consistent with this, exploratory measures of perceived warmth captured at the end of Studies 4A and 5 suggest that using affective language at the start and end made the agent seem warmer. Both warmth and competence perceptions were supported as mediators for our primary customer satisfaction outcome, and competence perceptions were supported for the secondary helpfulness outcome used by the firm in Study 1. See Web Appendix F for detail.

That said, measuring overall perceptions at the end of the interaction may not be the best approach to capturing what is going on. Temporal language effects may simply mean shifting the same amount of a feature (e.g., warmth) to a different moment, meaning that overall perceptions of warmth or competence might not always change. Consequently, future studies could use moment-to-moment measures (cf. Ramanathan and McGill 2007), to better investigate the mechanisms that underlie these temporal shifts. Future research could also consider more detailed measures of different dimensions of warmth (e.g., rapport-building versus empathetic).

Moderators also deserve further attention. To illustrate how one might approach such opportunities, ancillary analyses explored whether issue severity moderates the benefit of affective or cognitive language at particular conversational moments (see Web Appendix E, Study 6). Other situated aspects may also shape the effects. The best time to use affective language may be different in initial sales calls, for example, than when resolving existing customer issues. A single speaker monologue (e.g., voice actor in a radio ad), likely entails different temporal dynamics than two actors in dialogue. Results may also vary outside of traditional marketing contexts (e.g., doctor-patient conversations; Berger and Packard 2023). The importance of affective language may also be diminished when employees can build rapport using other means (e.g., facial expression).

Work could also explore conversational norms. While preferences for warmth and competence likely drive the observed effects, norms may also play a role. Customer service is a relatively constrained process (Marinova et al. 2018), which can lead to structured, ritualistic conversational norms (Goffman 1981) or expectations of how conversations will evolve. These structures are especially noticeable in early and late conversational moments known as "openings" and "closings" (Schegloff and Sacks 1973). Openings like "How are you today?" or "What can I do for you today?" are both normative for problem-solving conversations (Gafaranga and Britten 2005), but whether the warmer opening is just preferred or somehow violates the expected norms of service conversations is an open question. Future work should consider such possibilities, and whether the impact of violating other conversational norms (e.g., turn-taking, maxim violations; Grice 1975; Seedhouse 2005) may vary over conversational time.

Future work might also examine the role of culture. While warmth and competence are key dimensions across cultures, different cultures may have different values or baseline expectations around how much of each is desired. Spanish, Portuguese, and Italian people are seen as warmer, for example, while German and English people are seen as more competent (but less warm; Cuddy et al., 2009). Consequently, if they internalize these stereotypes, German and English consumers may prefer relatively more competence, for example, and less warmth.

The dynamic value of warmth and competence might also vary cross-culturally. Conversational norms differ across cultures (Kim 2017), so warmth may be less important at the beginning or end in some contexts. Even outside of culture, languages have different norms about when and how to express warmth and competence. Korean, for example, has a linguistic device that conveys warmth-related information at the end of most sentences (Lee and Ramsey 2000). In this language, limiting warmth to a conversation's start and end may be less beneficial, or difficult to achieve. Even within the same cultural context or language, variations in norms and expectations may shape what dynamic patterns are preferred. A conversation among Americans will often entail dyads from sub-cultures with different warmth and competence norms or stereotypes (e.g., southern vs. northeastern or Italian vs. Asian

Americans; Fiske 2018). Such cultural features, context (e.g., professional vs. personal), relative power, in- or out-group status, gender, and other factors likely shape conversation dynamics in complex ways. We hope future research may consider such potentially important variation.

Conclusion

This research begins to quantify when language matters. Beyond warmth and competence, the approach presented (and accessible for non-specialists at whenlanguagematters.net) should also be useful in studying advertising language, word of mouth, negotiation, message recall, and various other topics. We hope this work provides a useful framework for those examining conversations and other facets of human interactions.

REFERENCES

Aaker, Jennifer L., Emily N. Garbinsky, and Kathleen D. Vohs (2012), “Cultivating Admiration in Brands: Warmth, Competence, and Landing in the ‘Golden Quadrant’”, *Journal of Consumer Psychology*, 22, 191-194.

Abele, Andrea E. and Bogdan Wojciszke (2007), “Agency and Communion from the Perspective of Self Versus Other,” *Journal of Personality and Social Psychology*, 93(5), 751-763.

Aston, Guy (1995), “Say ‘Thank You’: Some Pragmatic Constraints in Conversational Closings,” *Applied Linguistics*, 16(1), 57-86.

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

Berger, Jonah and Grant Packard (2023), “Commentary: Using Language to Improve Health,” *Journal of Service Research*, 26(4), 514-516.

Berger, Jonah and Eric M. Schwartz (2011), “What Drives Immediate and Ongoing Word of Mouth?” *Journal of Marketing Research*, 48(5), 869-880.

Berger, Jonah, Stijn van Osselaer, and Chris Janiszewski (2023), “Variable Discovery and Prioritization: Casting a Wider Net,” *Working Paper*.

Berry, Diane S., James W. Pennebaker, Jennifer S. Mueller, and Wendy S. Hiller (1997), “Linguistic Bases of Social Perception,” *Personality and Social Psychology Bulletin*, 23(5), 526-537.

Blanding, Warren (1989), “Customer Service Communications: Telephone Systems and Procedures,” *Practical Handbook of Customer Service Operations*, 377-407.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan (2003), “Latent Dirichlet Allocation,” *Journal of Machine Learning Research*, 3, 993-1022.

Boersma, Paul and Vincent van Heuven (2001), "Speak and UnSpeak with PRAAT," *Glot International*, 5(9), 341-347.

Bolden, Galina B. (2008), "So What's Up?': Using the Discourse Marker So to Launch Conversational Business," *Research on Language and Social Interaction*, *41*(3), 302-337.

Bolton, Ruth N. and James H. Drew (1991), "A Longitudinal Analysis of The Impact of Service Changes on Customer Attitudes," *Journal of Marketing*, 55(1), 1-9.

Brody, Nathan (1994), "Cognitive Abilities," *Psychological Science*, 5(2), 63-68.

Brooks, Alison Wood, and Leslie John (2018), "The Surprising Power of Questions," *Harvard Business Review,* 96(3), 60-67.

Castleberry, Stephen B., C. David Shepherd, and Rick Ridnour (1999), "Effective Interpersonal Listening in The Personal Selling Environment: Conceptualization, Measurement, and Nomological Validity," *Journal of Marketing Theory and Practice*, 7(1), 30-38.

Cespedes, Frank V. and Christopher Wallace (2017), "Executives and Salespeople Are Misaligned – and The Effects Are Costly," *Harvard Business Review* (website).

Chang, Hsin Hsin, Yao-Chuan Tsai, Kit Hong Wong, Jian Wei Wang, and Fang Ju Cho (2015), "The Effects of Response Strategies and Severity of Failure on Consumer Attribution with regard to Negative Word-of-Mouth." *Decision Support Systems,* 71, 48-61.

Clark, Colin Mackinnon, Ulrike Marianne Murfett, Priscilla S. Rogers, and Soon Ang (2013), "Is empathy effective for customer service? Evidence from call center interactions." *Journal of Business and Technical Communication,* 27(2), 123-153.

Conlon, Donald E. and Noel M. Murray (1996), "Customer Perceptions of Corporate Responses to Product Comlaints: The Role of Explanations," *Academy of Management Journal*, 39(4), 1040-1056.

Cronin, Joseph J. and Steven A. Taylor (1992), “Measuring Service Quality: A Reexamination and Extension,” *Journal of Marketing*, 56(3), 55-68.

Cuddy, Amy J. C., Susan T. Fiske, Virginia Kwan, ,,, and Rene Ziegler (2009), “Stereotype content model across cultures: Towards universal similarities and some differences,” *British Journal of Social Psychology*, 48(1), 1-33.

Danescu-Niculescu-Mizil, Cristian, Robert West, Dan Jurafsky, Jure Leskovec, and Christopher Potts (2013), “No Country for Old Members: User Lifecycle and Linguistic Change in Online Communities,” *Proceedings of the 22nd International Conference on the World Wide Web*, 307-318.

de Ruyter, Ko and Martin G. M. Wetzels (2000), “The Impact of Perceived Listening in Voice-to-Voice Service Encounters,” *Journal of Service Research*, 2(3), 276-284.

Dixon, Matthew, Karen Freeman, and Nichlas Toman (2010), “Stop Trying to Delight Your Customers,” *Harvard Business Review*, July-August, 1-7.

Drollinger, Tanya, and Lucette B. Comer (1997), “A Conceptual Framework of Active Empathetic Listening in The Sales Encounter,” *National Conference on Sales Management Proceedings*, 142-145.

Drollinger, Tanya, Lucette B. Comer, and Patricia T. Warrington (2006), “Development and Validation of The Active Empathetic Listening Scale,” *Psychology & Marketing*, 23(2), 161-180.

Dubois, David, Derek D. Rucker, and Adam D. Galinsky (2016), “Dynamics of Communicator and Audience Power: The Persuasiveness of Competence and Warmth,” *Journal of Consumer Research*, 43(1), 68-85.

Durante, Federica, Coutrney Bearns Tablante, and Susan T. Fiske (2017), "Poor but Warm, Rich but Cold (and Competent): Social Classes in the Stereotype Content Model," *Journal of Social Issues,* 73(10, 138-157.

Fiske, Susan T., Amy J. C. Cuddy, and Peter Glick (2007), "Universal Dimensions of Social Cognition: Warmth and Competence," *Trends in Cognitive Sciences*, 11(2), 77-83.

Foutz, Natasha Z. and Wolfgang Jank (2010), "Prerelease Demand Forecasting for Motion Pictures Using Functional Shape Analysis of Virtual Stock Markets," *Marketing Science*, 29(3), 568-579.

Frank, Arthur W. (1982), "Improper Closings: The Art of Conversational Repudiation," *Human Studies*, 5, 357-370.

Gabor, Don (2011), *How to Start A Conversation and Make Friends*, New York, NY: Simon and Schuster.

Gafaranga, Joseph, and Nicky Britten (2005), "Talking an institution into being: The opening sequence in general practice consultations," *Applying Conversation Analysis*, 75-90.

Godfrey, Debra K., Edward E. Jones, Charles G. Lord (1986), "Self-Promotion Is Not Ingratiating," *Journal of Personality and Social Psychology*, 50(1), 106-115.

Goffman, Erving (1981), *Forms of Talk.* Philadelphia, PA: University of Pennsylvania Press.

Goodwin, Cathy and Ivan Ross (1992), "Consumer Responses to Service Failures: Influence of Procedural and Interactional Fairness Perceptions," *Journal of Business Research*, 25, 149-163.

Grewal, Dhruv, Dennis Herhausen, Stepan Ludwig, and Francisco Villarroel Ordenes (2021), "The Future of Digital Communication Research: Considering Dynamics and Multimodality," *Journal of Retailing*, in press.

Grice, Paul (1991). *Studies in the Way of Words*. Cambridge: Harvard University Press.

Grice, Paul (1975), "Logic and conversation," in *Speech Acts*, pp. 41-58, Leiden: Brill.

Gunturkun, Pascal, Till Haumann, and Sven Mikolon (2020), "Disentangling the Differential Roles of Warmth and Competence Judgments in Customer-Service Provider Relationships," *Journal of Service Research*, 23(4), 476-503.

Harris, Kim, and Steve Baron (2004), "Consumer-to-consumer conversations in service settings," *Journal of Service Research*, 6(3), 287-303.

Hayes, Andrew F. (2018), "Partial, Conditional, and Moderated Mediation: Quantification, Inference, and Interpretation," *Communication Monographs*, 85(1), 4-40.

Holoien, Deborah S. and Susan T. Fiske (2013), "Downplaying Positive Impressions: Compensation between Warmth and Competence in Impression Management," *Journal of Experimental Social Psychology*, 49(1), 33-41.

Huang, Karen, Michael Yeomans, Alison W. 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), 430-452.

Hui, Sam K., Tom Meyvis, and Henry Assael (2014), "Analyzing Moment-to-Moment Data Using A Bayesian Functional Linear Model: Application to TV Show Pilot Testing," *Marketing Science*, 33(2), 222-240.

Humphreys, Ashlee and Rebecca Jen-Hui Wang (2018), "Automated Text Analysis for Consumer Research," *Journal of Consumer Research*, 44, 1274-1306.

Jasmand, Claudia, Vera Blazevic, and Ko de Ruyter (2012), "Generating Sales While Providing Service: A Study of Customer Service Representatives' Ambidextrous Behavior," *Journal of Marketing*, 76(1), 20-37.

Judd, Charles M., Laurie James-Hawkins, Vincent Yzerbyt, and Yoshihisa Kashima (2005), “Fundamental Dimensions of Social Judgment: Understanding The Relations between Judgments of Competence and Warmth,” *Journal of Personality and Social Psychology*, 89(6), 899-913.

Kahn, Jeffrey H., Renee M. Tobin, Audra E. Massey, and Jennifer A. Anderson (2007), “Measuring Emotional Expression with the Linguistic Inquiry and Word Count,” *American Journal of Psychology*, 120, 263-286.

Kaski, Timo, Jarkko Niemi, and Ellen Pullins (2018), “Rapport Building in Authentic B2B Sales Interactions,” *Industrial Marketing Management*, 69, 235-252.

Keiningham, Timothy L., Bruce Cooil, Tor Wallin Andreassen, and Lerzan Aksoy (2007), “A Longitudinal Examination of Net Promoter and Firm Revenue Growth,” *Journal of Marketing*, 71, 39-51.

Kim, Min-Sun (2017, "Conversational Norms Across Cultures," *The International Encyclopedia of Intercultural Communication*, 1-9.

Kirmani, Amna and Margaret C. Campbell (2004), “Goal Seeker and Persuasion Sentry: How Consumer Targets Respond to Interpersonal Marketing Persuasion,” *Journal of Consumer Research*, 31(3), 573-582.

Kirmani, Amna, Rebecca Hamilton, Debora Thompson, and Shannon Lantzy (2017), “Doing Well vs. Doing Good: The Differential Effect of Underdog Positioning on Moral and Competent Service Providers,” *Journal of Marketing*, 81(1), 103-117.

Ko, Sei Jin, Charles M. Judd, and Diederik A. Stapel (2009), “Stereotyping Based on Voice in the Presence of Individuating Information: Vocal Feminity Affects Perceived Competence but Not Warmth,” *Personality and Social Psychology Bulletin*, 35(2), 198-211.

Kronrod, Ann, Amir Grinstein, and Luc Wathieu (2011), "Enjoy! Hedonic Consumption and Compliance with Assertive Messages," *Journal of Consumer Research*, 39 (June), 51-61.

Lee, Iksop and S. Robert Ramsey (2000), *The Korean Language*, New York, NY: SUNY Press.

Li, Xueni, Kimmy W. Chan, and Sara Kim (2019), "Service with Emoticons: How Customers Interpret Employee Use of Emoticons in Online Service Encounters," *Journal of Consumer Research*, 45(5), 973-987.

Luangrath, Andrea W., Joann Peck, and Victor A. Barger (2017), "Paralanguage and its Implications for Marketing Communications," *Journal of Consumer Psychology*, 27(1), 98-107.

Marinova, Detelina, Sunil Singh and Jagdip Singh (2018), "Frontline Problem-Solving Interactions: A Dynamic Analysis of Verbal and Nonverbal Cues," *Journal of Marketing Research*, 55(2), 178-192.

McBain, Jay (2020), "What I See Coming for The Channel in 2020," *Forrester Featured Blogs*.

McGuire, William J. (2000), "Standing on the Shoulders of Ancients: Consumer Research, Persuasion, and Figurative Language," *Journal of Consumer Research*, 27(1), 109-114.

Morgan, Blake (2017), "Customer Service is a $350 Billion Industry, and It's a Mess," *Forbes*.

Ng, Thomas W. H. and Daniel C. Feldman (2010), "Organizational Tenure and Job Performance," *Journal of Management*, 36(5), 1220-1250.

Ordenes, Franscisco V., Babis Theodoulidis, Jamie Burton, Thorsten Gruber, and Mohamed Zaki (2014), "Analyzing Customer Experience Feedback Using Text Mining: A Linguistics-Based Approach," *Journal of Service Research*, 17(3), 278-295.

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

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

Pallotti, Gabriele and Cecilia Varcasia (2008), "Service Telephone Call Openings: A Comparative Study on Five European Languages," *Journal of Intercultural Communication*, 17, 1-29.

Parasuraman, Ananthanarayanan, Leonard L. Berry, and Valarie A. Zeithaml (1991), "Understanding Customer Expectations of Service," *Sloan Management Review*, 32(3), 39-48.

Parasuraman, Ananthanarayanan and Valarie A. Zeithaml (2002), "Understanding and Improving Service Quality: A Literature Review and Research Agenda," *Handbook of marketing*, 339-367.

Parasuraman, Ananthanarayanan, Valarie A. Zeithaml, and Leonard L. Berry (1985), "A Conceptual Model of Service Quality and Its Implications for Future Research," *Journal of Marketing*, 49(4), 41-50.

Pennebaker, James W., Ryan L. Boyd, Kayla Jordan, and Kate Blackburn (2015), *The development and psychometric properties of LIWC2015*, https://repositories.lib.utexas.edu/bitstream/handle/2152/31333/LIWC2015_LanguageManual.pdf

Placencia, Maria E. (2004), "Rapport-Building Activities in Corner Shop Interactions," *Journal of Sociolinguistics*, 8(2), 215-245.

Pogacar, Ruth, Alican Mecit, Fei Gao, L. J. Shrum and Tina M. Lowrey (2022), "Language and Consumer Psychology," *APA Handbook of Consumer Psychology,* 451-470.

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

Radu, Alexandru Gabriel, Denni Arli, Jiraporn Surachartkumtonkun, Scott Weaven, and Owen Wright (2019), "Empathy and apology: The effectiveness of recovery strategies." *Marketing Intelligence & Planning*, 37(4), 358-371.

Ramanathan, Suresh and Ann L. McGill (2007), "Consuming with Others: Social Influences on Moment-to-Moment and Retrospective Evaluations of an Experience," *Journal of Consumer Research*, 34(4), 506-524.

Ramsay, James O. and Bernard W. Silverman (1997), *Functional Data Analysis*, New York: Springer.

Ramsey, Rosemary P. and Ravipreet S. Sohi (1997), "Listening to Your Customers: The Impact of Perceived Salesperson Listening Behavior on Relationship Outcomes," *Journal of the Academy of Marketing Science*, 25(2), 127-137.

Reece, Andrew, Gus Cooney, Peter Bull, Christine Chung, Bryn Dawson, Casey Fitzpatrick, Tamara Glazer, Dean Knox, Alex Biebscher, and Sebastian Marin (2022), "Advancing an Interdisciplinary Science of Conversation: Insights from a Large Multimodal Corpus of Human Speech," *PsyArXiv*, DOI: 10.31234/osf.io/ts43f

Rust, Roland T. and Tuck S. Chung (2006), "Marketing Models of Service and Relationships," *Marketing Science*, 25(6), 551-765.

Schegloff, Emanuel A. (1999), "Discourse, Pragmatics, Conversation, Analysis," *Discourse Studies*, 1(4), 405-435.

Schegloff, Emanuel A. and Harvey Sacks (1973), "Opening Up Closings", *Semiotica*, 8(4), 289-

327.

Schellekens, Gaby A. C., Peeter Verlegh, and Ale Smidts (2010), “Language Abstraction in Word of Mouth,” *Journal of Consumer Research*, 37 (2), 207–223.

Seedhouse, Paul (2005), "Conversation analysis as research methodology," in *Applying Conversation Analysis*, eds. Richards and Seedhouse, pp. 251-266, London: Palgrave Macmillan.

Singh, Sunil, Detelina Marinova, Jagdip Singh, and Kenneth Evans (2018), “Customer Query Handling in Sales Interactions,” *Journal of the Academy of Marketing Science*, 46, 837- 856.

Smith, Amy K., Ruth N. Bolton, and Janet Wagner (1999), “A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery,” *Journal of Marketing Research,* 36(3), 356-372.

Snyder, Hannah, Lars Witell, Anders Gustafsson, Paul Fombelle, and Per Kristensson (2016), “Identifying Categories of Service Innovation: A Review and Synthesis of The Literature,” *Journal of Business Research*, 69(7), 2401-2408.

Sood, Ashish, Gareth M. James, Gerard J. Tellis (2009), “Functional Regression: A New Model for Predicting Market Penetration of New Products,” *Marketing Science*, 28(1), 36-51.

Speckman, Paul (1988), “Kernel Smoothing in Partial Linear Models,” *Journal of the Royal Statistical Society, Series B*, 50(3), 413-436.

Spiro, Rosann L. and Barton A. Weitz (1990), “Adaptive Selling: Conceptualization, Measurement, and Nomological Validity,” *Journal of Marketing Research*, 27(1), 61-69.

Tausczik, Yla R. and James W. Pennebaker (2010), “The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods,” *Journal of Language and Social Psychology*, 29 (1), 24-54.

Van Zant, Alex B. and Jonah Berger (2020), “How The Voice Persuades,” *Journal of Personality and Social Psychology*, 118(4), 661-682.

Van Dolen, Willemijn M., Pratibha A. Dabholkar, and Ko de Ruyter (2007), “Satisfaction with Online Commercial Group Chat: The Influence of Perceived Technology Attributes, Chat Group Characteristics, and Advisor Communication Style,” *Journal of Retailing*, 83(3), 339-358.

Van Doorn, Jenny, Peter S. H. Leeflang, and Marleen Tijs (2013), “Satisfaction as A Predictor of Future Performance,” *International Journal of Research in Marketing*, 30(3), 314-318.

Wang, Jane-Ling, Jeng-Min Chiou, and Hans-Georg Muller (2016), “Functional Data Analysis,” *Annual Review of Statistics and Its Applications*, 3, 257-295.

Wojciszke, Bogdan and Andrea E. Abele (2008), “The Primacy of Communion over Agency and Its Reversals in Evaluations,” *European Journal of Social Psychology*, 39, 973-990.”

Wojciszke, Bogdan, Roza Bazinska, and Marcin Jaworski (1998), “On the Dominance of Moral Categories in Impression Formation,” *Personality and Social Psychology Bulletin*, 24(12), 1251-1263.

Xiong, Guiyang and Sundar Bharadwaj (2014), “Prerelease Buzz Evolution Patterns and New Product Performance,” *Marketing Science*, 33(3), 401-421.

Yao, Fang, Hans-Georg Muller, and Jane-Ling Wang (2005), “Functional Data Analysis for Sparse Longitudinal Data,” *Journal of the American Statistical Association*, 100(470), 577-590.

Yeomans, Michael, Maurice E. Schweitzer, and Alison Wood Brooks (2022), “The Conversational Circumplex: Identifying, Prioritizing, and Pursuing Informational and Relational Motives in Conversation,” *Current Opinion in Psychology*, 44, 293-302.

Zeithaml, Valarie A., Leonard L. Berry, and Ananthanarayanan Parasuraman (1996), "The Behavioral Consequences of Service Quality," *Journal of Marketing*, 60(2), 31-46.

Zhang, Justine, Sendhil Mullainathan, and Cristian Danescu-Niculescu-Mizil (2020), "Quantifying the Causal Effects of Conversational Tendencies," *Proceedings of the ACM on Human-Computer Interaction*, 4(131), 2-24.

Zhang, Qiang, Wenbo Wang, and Yuxin Chen (2020), "In-Consumption Social Listening with Moment-to-Moment Unstructured Data: The Case of Movie Appreciation and Live Comments," *Marketing Science*, 39(2), 285-457.

WEB APPENDIX A: STUDY 1 DESCRIPTIVES AND ADDITIONAL RESULTS

Figure A1: Study 1 Ratio of Agent Affective Language to Agent Cognitive Language (Affective / Cognitive Language)

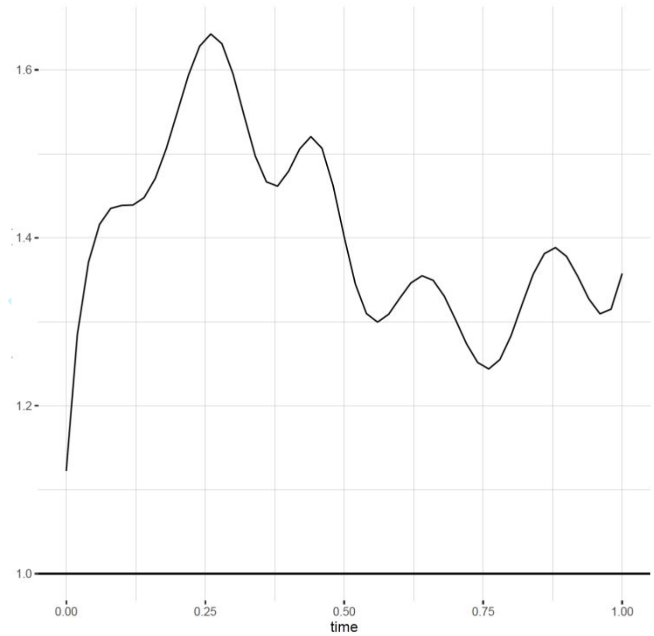
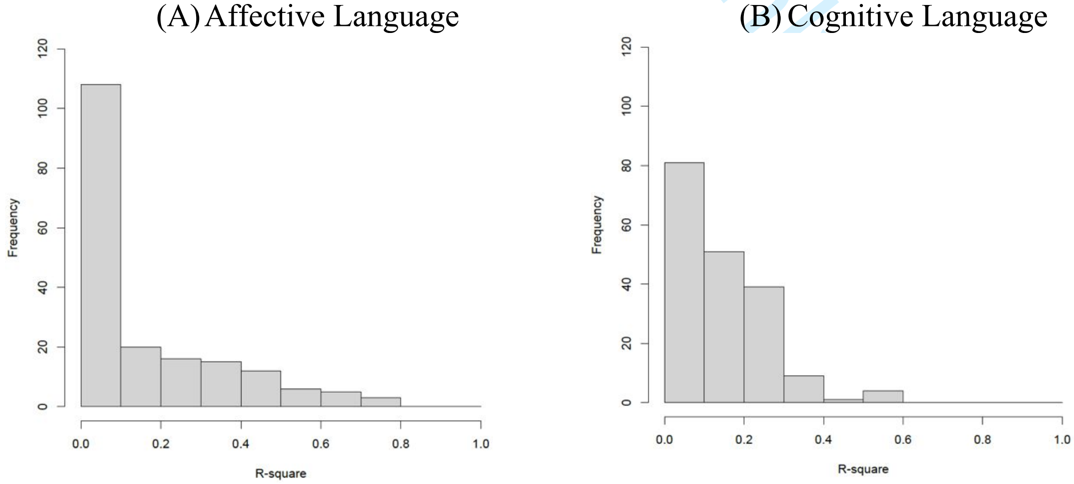


Figure A2: Study 1 Variance in Agent Language Explained by Customer Language (R^2)

(A) Affective Language

(B) Cognitive Language



Note: The histograms summarize the linguistic synchronicity of agent's and customer's affective and cognitive language across the 185 conversations. Overall, some level of conversational synchronicity happens more frequently for cognitive language, but synchronicity occurs more deeply for affective language in the fewer conversations in which it is present.

Figure A3: Study 1 Irregularity of Linguistic Features over Conversational Time (10 Sampled Calls)

(A) Agent Affective Language (B) Agent Cognitive Language

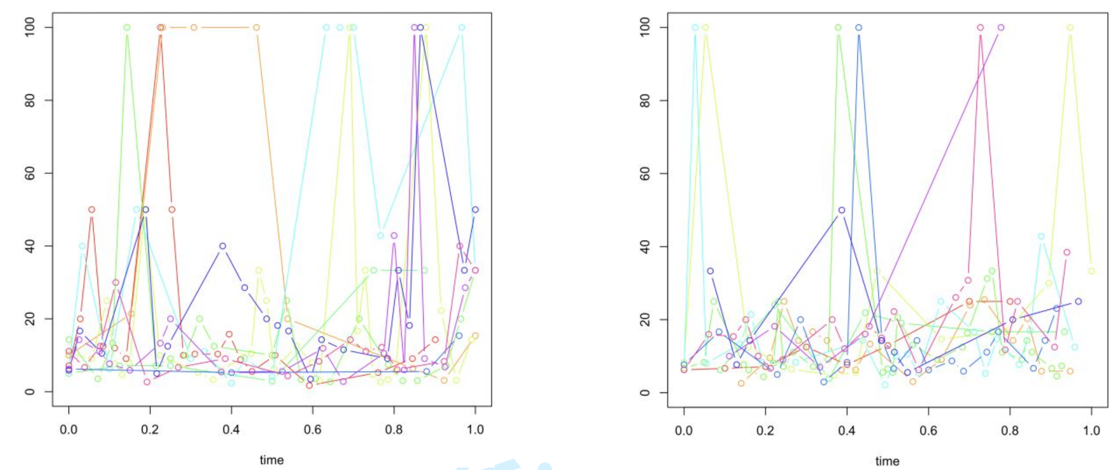


Figure A4: Study 1 Sparsity in Linguistic Measurements of Conversation

(A) Agent Affective Language (B) Agent Cognitive Language

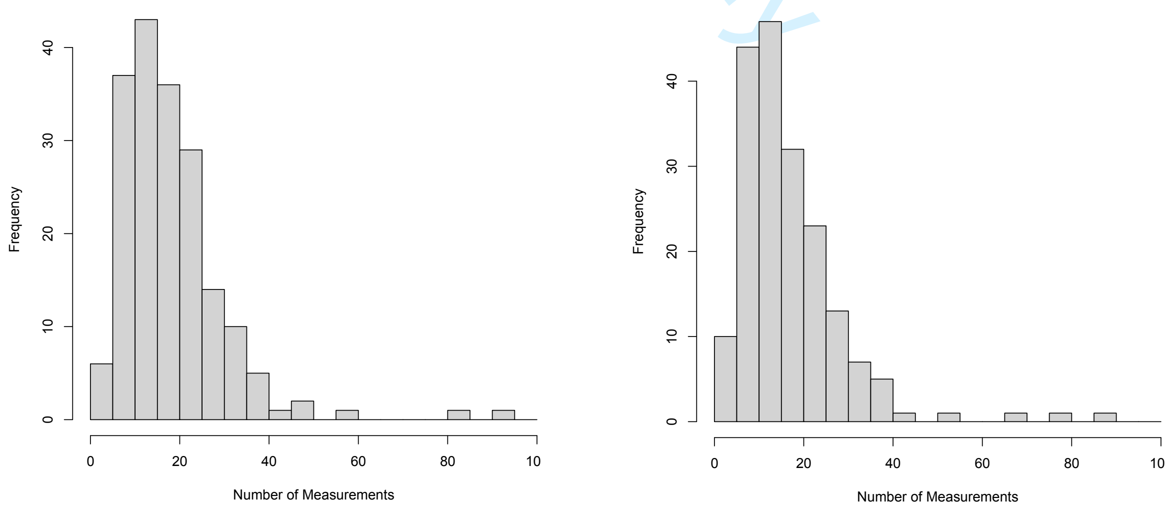


Table A1: Articles Analyzing Marketing Conversation Dynamics in Field Data

Article	Conversation Type	Product / Service	N (conversations)	N (unit of analysis)*	Number of controls**
Marinova, Singh, and Singh (2018)	Service	Airline	102	306	5
Singh, Marinova, and Singh (2020)	Sales	Heavy equipment	47	470	8
Singh, Marinova, Singh & Evans (2018)	Sales	Insurance	42	237	2
Current Research	Service	Retailer (S1)	185	12,410	45
		Airline (S2)	204	11,548	27
		Total	389	23,958	

*Unit of analysis for the respective field data studies is conversational stages (N = 3) in Marinova et al. (2018), conversational slices (N = 10) in Singh et al. (2020), time-ordered segments (M = 5.5) in Singh et al. (2018), and conversational turns (M = 66.8) in the current research.
**Current research Study 1 includes 22 traditional, structured data controls and 23 unstructured, wide data controls (total = 43). Study 2 includes 15 structured data controls, 12 unstructured data controls (total = 25).

Table A2: Study 1 Additional Call-Level Conversation Descriptives

	Mean	SD	Min	Median	Max
Number of words	1082.03	853.54	112.00	854.00	4385.00
Number of turns	66.75	44.49	13.00	60.00	337.00
Time per call (min.)	6.19	3.97	0.74	5.22	25.94

Table A3: Study 1 Summary Statistics

	Mean	SD	Min	Median	Max
Independent Measures					
Agent Affective Language	22.74	27.42	0.00	11.11	100.00
Agent Cognitive Language	16.03	14.79	0.00	12.50	100.00
Dependent Measures					
Customer Satisfaction	3.34	1.61	1.00	3.00	4.00
Orders 30 Days Post	0.76	1.76	0.00	0.00	23.00
Stuctured & Unstructured Controls					
Order	0.27	0.44	0.00	0.00	1.00
Shipping	0.27	0.45	0.00	0.00	1.00
Return	0.38	0.49	0.00	0.00	1.00
Product	0.05	0.22	0.00	0.00	1.00
Topic 1	0.09	0.06	0.02	0.07	0.41
Topic 2	0.07	0.06	0.01	0.04	0.35

Topic 3	0.08	0.05	0.02	0.07	0.45
Topic 4	0.08	0.07	0.02	0.07	0.60
Topic 5	0.07	0.06	0.01	0.05	0.45
Topic 6	0.09	0.10	0.02	0.06	0.61
Topic 7	0.07	0.06	0.01	0.05	0.44
Topic 8	0.08	0.05	0.01	0.07	0.28
Topic 9	0.07	0.05	0.01	0.06	0.30
Topic 10	0.07	0.04	0.01	0.06	0.38
Topic 11	0.08	0.05	0.02	0.06	0.28
Topic 12	0.08	0.07	0.02	0.06	0.58
Topic 13	0.09	0.05	0.01	0.07	0.29
Severity	2.61	0.94	1.00	2.50	5.00
Length	1082.03	853.54	112.00	854.00	4385.00
Resolved	0.80	0.40	0.00	1.00	1.00
Agent Tenure	412.38	650.85	0.00	216.00	3880.00
Agent Calls	4160.34	2456.80	37.00	4072.00	15010.00
Agent Female	0.61	0.49	0.00	1.00	1.00
Agent Social	12.35	16.85	0.00	8.57	100.00
Agent Perception	2.07	6.30	0.00	0.00	100.00
Agent Drive	6.48	10.73	0.00	0.00	100.00
Agent Time	17.10	15.06	0.00	17.39	100.00
Agent Informal	18.58	31.67	0.00	5.56	100.00
Agent Pitch	89.00	5.80	0.00	89.22	115.42
Agent Intensity	65.35	6.73	0.00	66.25	80.72
Customer Tenure	2177.19	1172.09	0.00	2123.00	4718.00
Customer LTV	6433.80	14600.02	68.00	2177.33	119762.85
Customer Region S	0.13	0.34	0.00	0.00	1.00
Customer Region E	0.36	0.48	0.00	0.00	1.00
Customer Region W	0.28	0.45	0.00	0.00	1.00
Customer Region MW	0.13	0.33	0.00	0.00	1.00
Customer Region OTHR	0.10	0.30	0.00	0.00	1.00
Customer Female	0.81	0.39	0.00	1.00	1.00
Att_Web	3.67	1.58	1.00	4.00	5.00
Att_Shop	3.47	1.71	1.00	4.00	5.00
Customer Affective Language	22.96	27.61	0.00	18.57	100.00
Customer Cognitive Language	21.51	19.79	0.00	16.67	100.00
Customer Social	7.88	16.00	0.00	0.00	100.00
Customer Perception	1.39	6.40	0.00	0.00	100.00
Customer Drive	4.85	13.28	0.00	0.00	100.00
Customer Time	14.79	17.16	0.00	12.50	100.00
Customer Informal	27.89	39.30	0.00	5.56	100.00
Customer Pitch	90.58	6.79	0.00	90.81	112.31
Customer Intensity	64.94	11.02	0.00	66.91	84.96
Orders 30 Days Pre	1.30	1.71	0.00	1.00	18.00

Table A4: Variance Inflation Factors

A variable inflation factor (VIF) is used in ordinary least squares regression to quantify the severity of multicollinearity (James et al. 2017). How to define VIF in a functional setting is less clear, as it involves regressions from scalar to functional variables (e.g., DV is a scalar while IVs are trajectories). To compute a "functional VIF" we use a functional-to-function regression to report the functional quasi R^2 among the four primary functional variables – agent's warm and competent language as well as customer's affective and cognitive language. We then perform standard VIF calculations for the traditional structured controls. The table below reports the results. VIF values are fairly low, indicating the multicollinearity among the variables is not severe (considerably below the standard VIF cutoff of 10, or the more conservative cutoff of 5).

Variable	VIF*
Focal Functional Variables	
Affect_A	1.66
Cognition_A	1.96
Affect_C	1.26
Cognition_C	1.58
Structured Controls	
Severity	1.61
Length	1.96
Resolved	1.50
Return	4.79
Order	4.18
Shipping	4.56
Product	--
Agent Tenure	1.34
Agent Calls	1.23
Agent Female	1.28
Customer Tenure	1.39
Customer LTV	1.30
Customer Female	1.16
Customer Reg N	--
Customer Reg W	3.50
Customer Reg E	3.97
Customer Reg S	2.69
Customer Reg MW	2.59
Attitude Web	2.48
Attitude Shop	2.78

**Quasi R^2 for functional variables*
Note: Product and Customer Region N are constants for the Call Category and Region dummies, respectively.

Table A5: Study 1 Call-Level Linear Regression for Customer Satisfaction after Lasso

	Estimate	SE	p-stat
(Intercept)	0.50	0.47	0.29
Agent Affective Language	0.05	0.03	0.04
Agent Cognitive Language	-0.04	0.03	0.15
Topic 1	2.67	1.34	0.05
Topic 2	-3.99	1.17	0.00
Topic 7	-2.40	1.08	0.03
Cust. Region MW	-0.37	0.18	0.04
Att_Web	0.24	0.06	0.00
Att_Shop	0.46	0.05	0.00
Cust. Perception	0.12	0.05	0.01
Cust. Informal	0.03	0.01	0.05

Table A6: Study 1 Call-Level Poisson Regression for Customer Purchases after Lasso

	Estimate	SE	p-stat
(Intercept)	-0.07	0.47	0.89
Agent Affective Language	-0.08	0.04	0.05
Agent Cognitive Language	-0.02	0.04	0.54
Orders 30 Pre	0.21	0.01	0.00

Table A7: Study 1 Agent Language and Customer Satisfaction

	Model 1	**Model 2**	**Model 3**
Conceptual Predictors			
Affect_A_1	0.015 (0.014)	0.004 (0.015)	0.002 (0.009)
Affect_A_2	0.037 (0.022)^	0.029 (0.012)*	0.018 (0.009)*
Affect_A_3	0.059 (0.022)**	0.068 (0.021)**	0.063 (0.013)***
Affect_A_4	-0.014 (0.019)	-0.035 (0.02)^	0.006 (0.002)**
Affect_A_5	-0.100 (0.072)	-0.042 (0.069)	0.037 (0.048)
Affect_A_6	0.086 (0.232)	-0.037 (0.212)	-0.016 (0.129)
Cognition_A_1	-0.007 (0.003)*	-0.023 (0.010)**	-0.033 (0.014)**
Cognition_A_2	0.015 (0.007)*	0.042 (0.025)^	0.014 (0.008)^
Cognition_A_3	0.072 (0.041)^	0.071 (0.036)^	0.019 (0.010)^
Cognition_A_4	-0.095 (0.069)	-0.073 (0.062)	-0.051 (0.030)^
Cognition_A_5	0.345 (0.337)	-0.019 (0.307)	-0.031 (0.185)
Cognition_A_6	2.607 (2.333)	0.050 (1.222)	0.550 (0.802)
Structured Controls			
Severity		-0.129 (0.049)**	-0.089 (0.042)*
Length		0.000 (0.000)	0.000 (0.000)
Resolved		0.372 (0.224)^	0.034 (0.023)
Return		0.435 (0.432)	0.604 (0.258)
Order		0.543 (0.437)	0.521 (0.280)
Shipping		0.389 (0.447)	0.373 (0.275)
Product		--	--
Agent Tenure		0.000 (0.000)*	0.000 (0.000)^
Agent Calls		0.000 (0.000)	0.000 (0.000)

Agent Female		-0.128 (0.221)	-0.055 (0.139)
Agent Affect Language		0.004 (0.007)	0.005 (0.007)
Agent Cognition Language		0.003 (0.005)	0.003 (0.006)
Customer Tenure		0.000 (0.000)	0.000 (0.000)
Customer LTV		0.000 (0.000)	0.000 (0.000)
Customer Female		-0.013 (0.257)	-0.123 (0.155)
Customer Reg N		--	--
Customer Reg W		0.622 (0.559)	0.576 (0.584)
Customer Reg E		0.9 (0.942)	0.908 (0.499)^
Customer Reg S		0.667 (0.578)	0.664 (0.586)
Customer Reg MW		0.414 (0.572)	0.422 (0.580)
Attitude Web		0.205 (0.055)***	0.115 (0.042)**
Attitude Shop		0.443 (0.054)***	0.473 (0.116)***
Unstructured Controls			
Drives_A_1			-0.035 (0.023)
Drives_A_2			0.033 (0.026)
Drives_A_3			0.021 (0.030)
Drives_A_4			0.120 (0.064)^
Drives_A_5			-0.866 (0.810)
Drives_A_6			4.93 (9.478)
Pitch_A_1			0.02 (0.027)
Pitch_A_2			0.419 (0.357)
Pitch_A_3			-1.527 (0.916)^
Pitch_A_4			-2.352 (1.884)
Pitch_A_5			0.952 (3.28)
Intensity_C_1			-0.021 (0.014)
Intensity_C_2			0.041 (0.033)
Intensity_C_3			0.065 (0.052)
Intensity_C_4			0.144 (0.115)
Intensity_C_5			0.166 (0.130)
Intensity_C_6			1.806 (0.978)^
Intensity_C_7			1.522 (0.983)
Topic 1			1.234 (2.174)
Topic 2			-5.975 (3.568)^
Topic 4			-2.726 (1.967)
Topic 9			0.624 (1.796)
Synchronicity			0.123 (0.074)^
Intercept	3.096 (0.122)***	3.504 (0.727)***	1.806 (0.613)**

*^ p<0.1   * p<0.05   ** p<0.01   *** p<0.001*

Note: For the dynamic predictors and control variables, the word represents the language variable (e.g. "affect"), the letter the speaker (Agent (A) or Customer (C)), and the number indicates the eigen component number (e.g,. Affect_A_1). Product and Customer Region N are constants for the Call Category and Region dummies, respectively.

WEB APPENDIX B: STUDY 1 WIDE DATA CONTROL DETAILS, ROBUSTNESS TESTS, BENCHMARKS, AND SIMULATIONS

Model Approach for Additional Unstructured Controls

As discussed, we move beyond prior research by attempting to consider an even larger set of dynamic linguistic and paralinguistic features available in conversation's wide data. As a result, in Equation (8), the total number of wide data controls available to consider $(L + J)$ becomes comparable to the number of observations, therefore the model is likely to overfit, resulting in less meaningful model estimates. To address this possibility, the regression needs to be regularized so that the additional wide data controls that might also be contributors to the relationship of conceptual and substantive interest (the warmth/competence trade-off in customer service) can be automatically selected for efficient model inference.

Conventional variable selection methods such as stepwise regression (e.g., Foutz and Jank 2010) are not appropriate for two reasons. First, solutions from stepwise regression are path-dependent as the approach is a *greedy* algorithm that finds local optima in every step, but often fails to reach generally optimal variable selection (lack of oracle properties; Zou 2006). Second, stepwise regression does not allow *group-wise* variable selection, whereas the selection of additional wide data functional controls corresponds to selecting from the L groups of eigen scores in (8). That is, for a given functional control $X_l(t)$, there are two possible scenarios: either all the $\{b_{li}\}_{i=1}^{I_l}$ are completely eliminated, or all of them are chosen to be included in the regression.

To overcome these challenges, we utilize Group-Lasso regularization (Yuan and Lin 2006; Meier, van de Geer, and Bulhmann 2008; Yang and Zou 2015) to avoid path-dependency and to retain the functional control variable groupings (i.e., feature specific sets of eigen-components) after selection. This approach retains the wide data dynamic language or paralanguage controls that aid model fit, and penalizes out of the model those additional wide data dynamic controls that either do not enhance model fit, or contribute to prediction error.

The shrinkage and variable selection method, Lasso (Tibshirani 1996), has been widely applied in statistics and machine learning for high dimensional data analysis. Yuan and Lin (2006) proposed a generalization of Lasso for group-wise variable selection and regularization. To answer our central research questions around the dynamics of the well-established importance of affective and/or cognitive language in customer service and more broadly (e.g., Holoien and Fiske 2013; Kirmani et al. 2017; Marinova et al. 2018), we keep the two functional predictors unpenalized in the L1 regularization procedure (Chen et al. 2016; Heinze, Wallisch, and Dunkler 2018). Assuming the wide data controls in our model can be divided into D non-overlapping groups, where D is determined by the number of controls and the truncation of eigen components for each functional variable, Group-Lasso attempts to minimize

$$\frac{1}{2}\left\| g(E[y]) - \alpha_a - \vec{b}_A\vec{\omega}_A - \vec{b}_C\vec{\omega}_C - \sum_{d=1}^{D} \vec{b}_d\vec{\omega}_d \right\|_2^2 + \lambda \sum_{d=1}^{D} \sqrt{\dim\left(\vec{b}_d\right)} \left\|\vec{b}_d\right\|_2 \tag{9}$$

where subscripts “A” and “C” denote the affective and cognitive language components respectively. The Group-Lasso procedure suppresses a subset of groups of coefficients to zero to encourage a simpler and more efficient generalized linear model. Solving the above penalized least squares is computationally costly, so we follow Yang and Zou (2015) and implement the groupwise-majorization-descent (GMD) algorithm to achieve fast computation of Group-Lasso for the simultaneous selection of functional and scalar variables. To determine the optimal value of penalty parameter λ, we first calculate the maximum penalty parameter λ_{max} such that none of the penalized groups are active in the model. Then we construct a multiplicatively decaying grid for possible λ values starting at λ_{max}, and use leave-one-out cross-validation to pick the best penalty parameter from the grid. The resultant λ values are [211.43, 105.71, 52.85, 26.42, 13.21, 6.60, 3.30, 1.65, 0.82, 0.41, 0.20].

Alternative Measures of Affective and Cognitive Language Styles

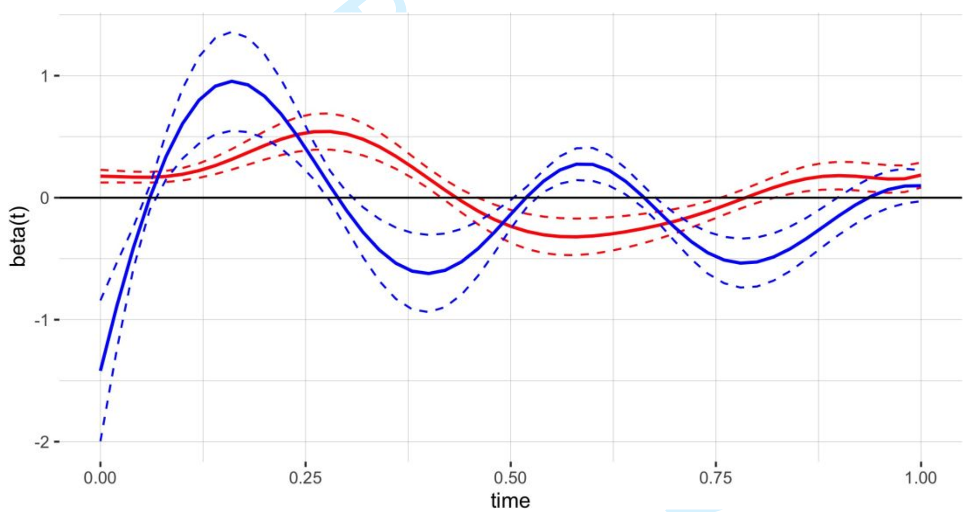
The affective and cognitive language measures used here have been extensively validated in prior work (cf. reviews by Kahn et al. 2007 and Tausczik and Pennebaker 2010), but one could wonder whether they might miss certain idiosyncratic features of customer service conversation. To address this possibility, we apply custom dictionaries from prior service research (Marinova et al. 2018; Singh et al. 2018). These works combined established dictionaries (LIWC) and human judging to develop custom lists of "relating" (i.e., affective) words ($N = 247$) and "resolving" (i.e., cognitive) words ($N = 649$). We scored all agent and customer conversational turns using this approach, and estimated our main model with these alternative measures.

Results are similar. As before, customers are more satisfied when agents use the alternative affective language measure ("relating") during a conversation's start and end, but less satisfied when this language is used in the middle (Figure B1). Similarly, for cognitive language, customers are more satisfied when agents use "resolving" language in the middle of the call, but less satisfied when such language occurs at the beginning.

Note that these results differ from prior work. The research that developed these dictionaries (Marinova et al. 2018) observed that *only* agent cognitive language was positively linked to their dependent measure (i.e., human judgement of customer emotion). They found that affective language *impeded* cognitive language's benefits when both were included in the model, supporting the warmth/competence trade-off and a recommendation to focus exclusively on competence-oriented cognitive language. These differences are likely driven by our dynamic modeling approach, but may also be due in part to distinctions in the specific customer service contexts (airline counter service vs. online fashion retailer), or different dependent measures (e.g., third-party judgment of displayed affect vs. customer satisfaction self-reports).

Figure B1: Study 1 Beta Curves for Agent "Relating" and "Resolving" (B) Language

Model 3 (Model 2 + unstructured controls after Group-Lasso)



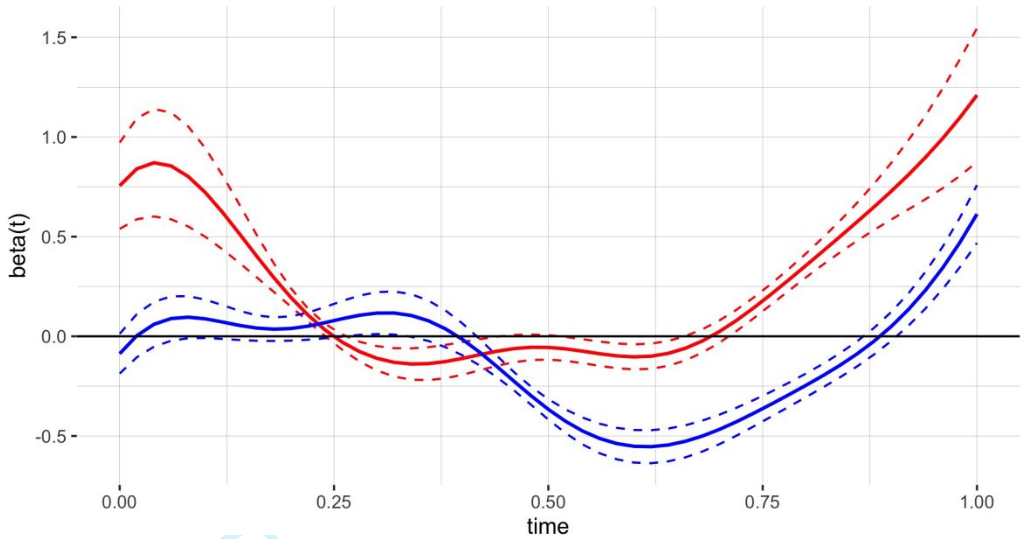
Red line: Relating Language; Blue line: Resolving Language
Dotted lines: pointwise 95% confidence intervals

Valenced Subsets of Affective Language

While LIWC's affective process dictionary is often used to capture warmth, one could argue that "warm" affective language should contain only positive emotional words (e.g., happy and wonderful) and exclude negative ones (e.g., sad and disappointed). Agents often use negative affective language in a warm manner to convey empathy (e.g., "I'm disappointed we didn't deliver your order on time"), but to test the contribution of each valence we repeat the main analysis incorporating agents' positive and negative affective words as separate predictors using the LIWC posemo and negemo dictionaries that, when combined, represent the affective processes dictionary used for the main analysis.

Results are again similar. The beta curve for positive affective language is close to that of the full affective language dictionary, while negative affective language also appears to contribute positively, albeit only at the end (Figure B2). A review of the negative affect words used in the conversational closings reveals that the presence of words like "sorry," "problem," and "wrong" are positively correlated with customer satisfaction (i.e., "Sorry about that" or "Glad we could fix the problem"). Our approach appears to capture such subtle conversational language features well. Positive affective language has similar relationships, but the effects are reduced for negative affective language, suggesting that the benefit of warmth arises primarily from positive affective language.

Figure B2: Study 1 Beta Curves for Agent Positive Affective and Negative Affective Language in Relation to Customer Satisfaction



Red line: Positive Affect Language; Blue line: Negative Affect Language
Dotted lines: pointwise 95% confidence intervals

Purchase Behavior Outcome

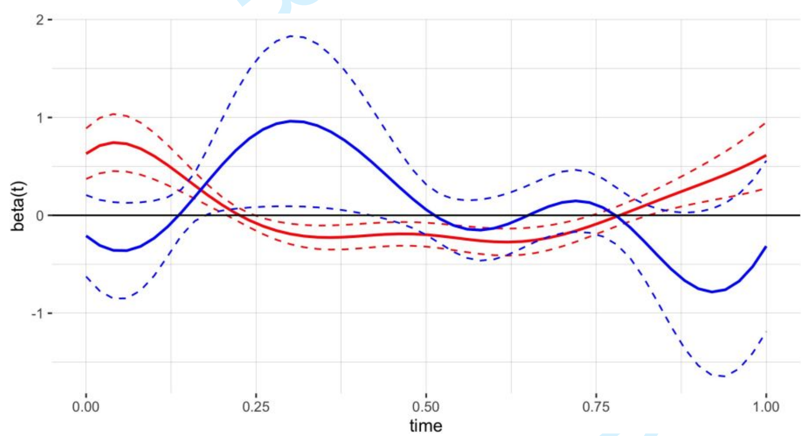
While we are primarily focused on customer satisfaction, one might wonder whether they are robust when extended to a behavioral measure like purchase. To consider this possibility, we apply a functional Poisson regression with a Log link function in (8) to estimate how agent affective and cognitive language relate to downstream purchase behavior (i.e., order count). The Poisson model introduces the same sequence of control variables as the functional linear regression. Importantly, to account for customer variation in baseline purchase behavior, the model further includes the number of orders each customer placed up to 30 days prior to the conversation (*Orders 30 Pre*).

Even examining this more behavioral measure, results remain similar (Figure B3). Customers purchased more when agents use affective language at the beginning and end of the call, but cognitive language in the middle. The only noteable difference for the model considering robustness to downstream purchase behavior is affective language at the end. While this was significant and positive for affective language in Model 1, the importance of affective language at the end of the conversation was non-significant when the 20 structured controls were added (Model 2).

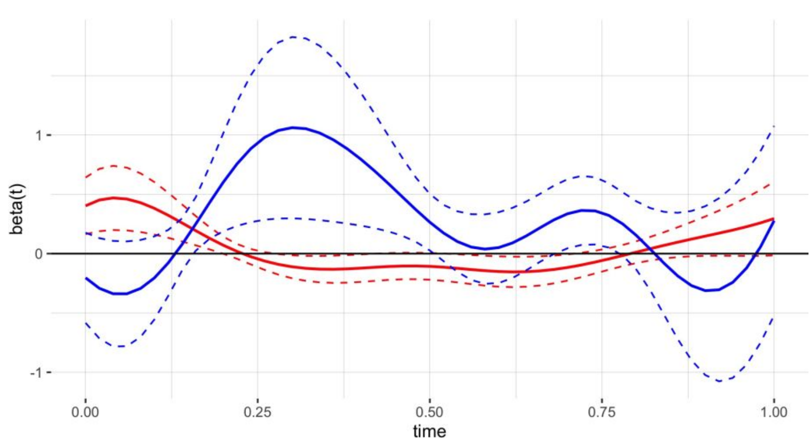
Notably, the model that attempts to add the additional 100+ unstructured control paramaters (Model 3) failed to converge due to a matrix inversion error. This likely occurs because such data is too wide for a Poisson model.

Figure B3: Study 1 Agent Language and Downstream Purchase Behavior

Model 1 (no controls)



Model 2 (Model 1 + controls)



Red line: Affective Language; Blue line: Cognitive Language
Dotted lines: pointwise 95% confidence intervals

Benchmark Model Comparisons

Although our approach uniquely produces moment-to-moment insights over the course of a conversation, one may wonder whether it performs better than competing benchmarks. To test this, we compare our full model (Model 3) against several benchmark models (BM1, BM2, …).

BM1: Simple "what" analysis. To examine the standard approach used in most prior work, this model aggregates all turns together, assessing agent affective and cognitive language over the course of the conversation in a multivariate Lasso regression. It includes all static controls, and conversation-level averages of the dynamic language and paralanguage features.[1]

BM2: "What" analysis with conversational stages. Following Marinova, Singh, and Singh (2018), judges dummy coded each conversational turn into one of three fixed stages: *Sensing*, *Seeking*, and *Settling*. The *Sensing* stage in that research averaged 12% of the interaction, the *Seeking* stage about 83%, and the *Settling* stage the last 5% of a given

[1] Model estimates suggest that if we had only analyzed these language features at conversation-level, consistent with prior research, we would have concluded agents should use only one of either affective or cognitive language, but not both (Web Appendix Tables A5 and A6). The call-level model estimates indicate customer satisfaction has a positive relationship with agent affective language ($b = 0.05$, $p < 0.05$), and a non-significant relationship with agent cognitive language ($b = -0.04$, $p > 0.1$). These findings are more consistent with the psychology literature's suggestions of prioritizing warmth than the competence-oriented speaking style recommended in recent customer service research.

conversation. In each, we compute the agent's use of affective and cognitive language. We also include the same set of controls as in Model 1.

BM3: Functional model with all wide data. Here we simply use all the wide data controls we obtained to estimate the model specified in (8), without consideration of model overfitting.

BM4: Homogeneous Functional model with Group-Lasso. In this model we integrate Group-Lasso into BM3 to avoid overfitting due to the additional wide data controls, but ignore the agent heterogeneous effect.

Results. Table B1 reports model comparisons. When conducting out-of-sample prediction, we hold out conversations one by one using leave-one-out cross-validation (Hui et al. 2014).[2]

Traditional "what" analyses (BM1 and BM2) that do not account for moment-to-moment conversation dynamics yield poorer in-sample and out-of-sample predictions than our functional framework (BM4 and our model). The functional regression model that uses high dimensional data (BM3) improves in-sample fit relative to its counterparts with Group-Lasso (BM4 and our model), but its out-of-sample prediction deteriorates significantly due to overfitting. One can discern that the out-of-sample prediction of BM3 is sometimes even worse than the static "what" analyses (BM1 and BM2), highlighting the importance of model regularization in functional regression on wide data (Model 3). Further, the incorporation of the heterogeneous agent intercept offers little benefit (our model vs. BM4), likely because the number of conversations (N = 185) is close to the number of agents (N = 130). Taken together, this indicates our approach offers superior predictive performance relative to previous models.[3]

---

[2] The smoothing for the functional variables is done separately for each training dataset.
[3] In addition to the benchmark model comparisons, one could still wonder whether prior work's suggestion to exclusively use an affective or cognitive style may be best, or how much "when" one uses these styles matters if one tries to use both. To probe

Table B1: Study 1 In-Sample Fit and Out-of-Sample Prediction

		BM1	BM2	BM3	BM4	Our Model 3
In-Sample Fit	RMSE	1.52	1.45	0.52	0.62	0.62
	MAD	1.35	1.22	0.37	0.44	0.43
	Correlation*	0.30	0.21	0.96	0.94	0.94
Out-of-Sample Prediction	RMSE	1.65	1.63	2.28	0.99	0.98
	MAD	1.49	1.39	1.73	0.78	0.78
	Correlation*	0.23	0.31	0.41	0.82	0.82

**Correlation between the predicted value and the actual outcome.*

Simulations

Model comparisons presented in the main paper suggest our approach to capturing conversational dynamics enhances the predictive benefit of understanding when affective or cognitive language is beneficial. But one might still wonder how the model's dynamic recommendations (i.e., using more affective language at start and end, and cognitive language in the middle) should perform relative to the exclusively cognitive or affective approaches recommended in prior research. Similarly, one could ask how much *when* one uses each of these language styles matters if both affective and cognitive language are used in a single interaction.

To begin to answer these questions, we performed a series of simulations. Because our model identifies when affective and cognitive language should be used, but not the optimal level of these features at a given moment, the simulations utilize the average observed levels of affective and/or cognitive language at each conversational moment, and then turn that language feature "on" or "off" at different moments based on the simulation condition. These simulations

these questions, we performed a series of simulations comparing our model with various alternatives. Results further support our dynamic model approach. See Web Appendix.

compare alternative approaches to the dynamic language use suggested by our modeling estimates. Consequently, the simulated improvements in satisfaction and purchases should be considered cautiously as optimistic ceilings rather than expected outcomes because they assume that agents are able to perfectly follow the moment-moment optimal timing of affective and cognitive language. What's more, the model allows for predicted outcomes outside the bounds of the customer satisfaction measure used by the firm.

First, we compare our model's recommended approach to the marketing literature's recommendation to be competence-oriented throughout the interaction. The simulation suggests that, at the optimistic ceiling, employees who follow the timing of affective and cognitive language suggested in the current approach (Figures 2 and 3) would see a 2.50 point increase in customer satisfaction ($p < 0.01$) and 3.42 more purchases in the 30 days following the call ($p < 0.01$) over this simulated competence-only baseline. For a more conservative test, we also compare our approach to a competence-only approach that uses cognitive language only, but emphasized at the "right times" (per Figures 2 and 3). Results further support the notion that using both affective and cognitive language at the right times, rather than only cognitive language at the right times, should have beneficial effects, i.e., difference in customer satisfaction $= 2.06$ ($p < 0.01$) and in purchases $= 2.84$ ($p < 0.01$).

Results are similar when we compare the current approach to the psychology literature's suggestion to be affective (or warm) throughout the interaction, i.e., difference in customer satisfaction $= 2.42$ ($p < 0.01$) and in purchases $= 3.69$ ($p < 0.01$). A comparison to being affective only but at the "right times" shows similar results, i.e., difference in customer satisfaction $= 1.36$ ($p < 0.01$) and in purchases $= 1.87$ ($p < 0.01$).

Second, we consider a comparison which acknowledges that affective and cognitive language can fruitfully co-exist in a single interaction but ignores the possibility that *when* these speaking styles are used matters. To do so, we simulate a scenario in which the two speaking styles are turned on at the mean observed level at every point in conversational time. Speaking both affective and cognitively at the "right times" rather than at all times results in a simulated improvement of 1.49 points in customer satisfaction ($p < 0.05$) and an incremental 2.93 purchases in the 30 days after the call ($p < 0.05$).

Taken together, while the size of the simulation results should be considered cautiously as optimistic ceilings rather than expected values, they support the benefits of using *both* affective and cognitive language rather than only one, and of considering *when* to use each of these approaches over the course of a conversation rather than merely *what* (i.e., more affective or cognitive language overall).

WEB APPENDIX C: STUDY 2 DESCRIPTIVES AND ADDITIONAL RESULTS

Descriptives

Figure C1: Study 2 Means and 95% Confidence Bands of Focal Features over Conversational Time for the Airline Dataset

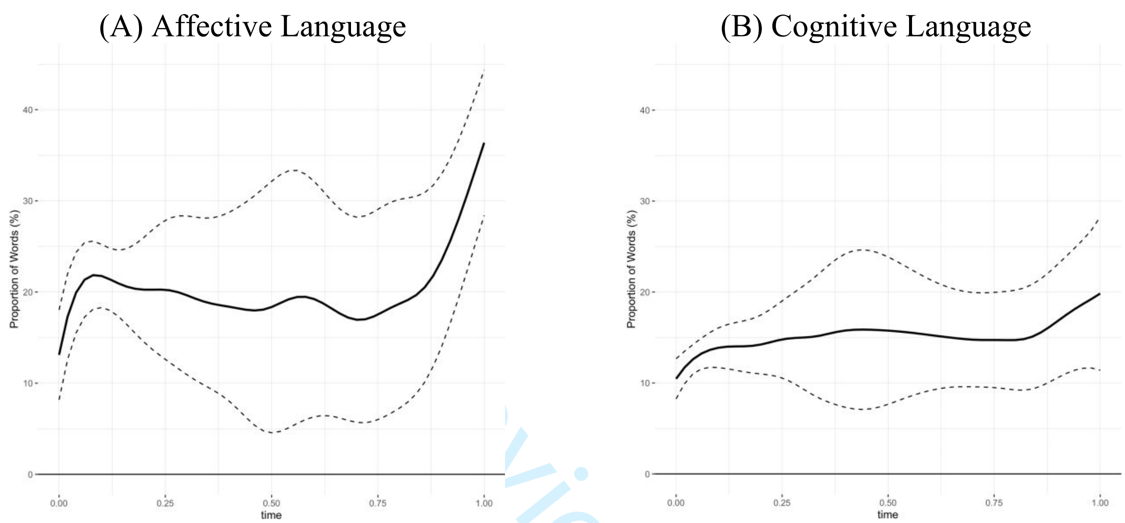


Table C1: Study 2 Parameter Estimates for Willingness to Recommend

	Model 1	Model 2	Model 3
<u>Conceptual Predictors</u>			
Affect A_1	0.357 (0.120)**	0.338 (0.110)**	0.349 (0.117)**
Affect_A_2	0.216 (0.500)	0.204 (0.115)$^{\wedge}$	-0.007 (0.004)$^{\wedge}$
Affect_A_3	-0.432 (0.226)$^{\wedge}$	-0.347 (0.647)	-0.311 (0.674)
Cognition_A_1	-0.199 (0.120)$^{\wedge}$	-0.245 (0.130)$^{\wedge}$	-0.123 (0.057)*
Cognition_A_2	0.515 (0.177)**	0.547 (0.250)**	0.320 (0.161)*
Cognition_A_3	0.976 (1.545)	1.162 (1.419)	-0.692 (1.527)
Cognition_A_4	2.639 (6.018)	4.245 (5.577)	4.442 (5.665)
<u>Structured Controls</u>			
Length		0.000 (0.001)	0.001 (0.001)
Exchange		0.818 (0.498)	1.038 (0.552)$^{\wedge}$
Refund		NA	NA
Agent Affect Language		0.003 (0.002)	0.002 (0.002)
Agent Cognition Language		0.002 (0.002)	0.002 (0.003)
Department 1		2.814 (0.885)**	1.619 (0.918)$^{\wedge}$
Department 2		4.195 (1.225)**	4.044 (1.253)**
Department 3		-0.703 (2.426)	-2.648 (2.431)
Department 4		NA	NA
Category 1		2.453 (0.533)***	1.777 (0.581)**
Category 2		0.148 (1.853)	-0.602 (1.81)
Category 3		2.334 (1.512)	2.418 (1.501)
Category 4		3.471 (1.640)*	2.858 (1.629)$^{\wedge}$

Category 5		--	--
Category 6		NA	NA
Unstructured Controls			
Affect_C_1			0.047 (0.041)
Affect_C_2			-0.034 (0.013)*
Affect_C_3			0.151 (0.181)
Affect_C_4			1.040 (0.777)
Time_A_1			0.616 (0.225)**
Time_A_2			-0.073 (0.728)
Time_A_3			-1.111 (1.428)
Time_A_4			-1.742 (1.692)
Topic 1			-3.158 (1.817)^
Topic 8			3.623 (5.773)
Topic 9			3.016 (5.739)
Topic 12			1.307 (0.782)^
Intercept	5.912 (0.244)***	1.308 (0.381)**	-3.453 (0.972)***

*^ p<0.1   * p<0.05   ** p<0.01   *** p<0.001*

Note: For dynamic predictors and controls, the word represents the language variable (e.g. "affect"), the letter represents the speaker (Agent (A) or Customer (C)), and the number indicates the eigen component number. NAs are identification level for dummy sets. Call category 5 was empty (not observed in the sample provided).

Pooling the Study 1 and Study 2 Datasets[4]

One might wonder how the field data results would look if the retailer (Study 1) and airline (Study 2) datasets were combined and examined in a single model. There are several issues, however, that make such an analysis problematic. First, the two firms provided different outcome measures obtained on distinct scales. Second, before the regression, we must first run functional Karhunen-Loève expansion to obtain eigen components. When we apply the functional expansion to the pooled data, it is impossible to include the fixed effect indicating observation-dataset affiliation in the expansion process. Alternatively, if we run the functional expansion separately on the two datasets, then not only do we disallow information sharing, but we also end up with two sets of distinct eigen components that are difficult to combine (e.g., it is

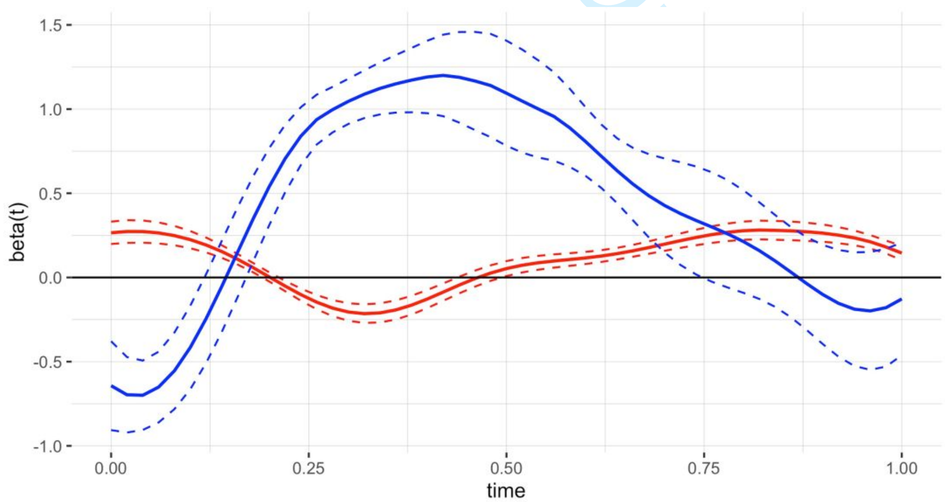
[4] We thank a reviewer for suggesting this pooling data exercise.

not clear how to pool together two eigen functions of different sizes for the same language feature).

That said, to explore data pooling, we neglect the conceptual underpinnings and standardize the two dependent measures to have comparable scales. Then we pool the independent measures common to both datasets and apply the functional orthogonal expansion to the pooled data. In the subsequent functional regression, we include a fixed effect to indicate the observation - dataset affiliation. Note that the main predictors in the functional regression are eigen scores computed from the pooled Karhunen-Loève expansion, thus the estimated fixed effect is unable to tease out the effect of data pooling as in ordinary regression scenarios.

Results are similar (see Figure C2). As noted, however, these results should be interpreted with caution given the fixed effect in the final regression may not resolve the inconsistencies between the two datasets.

Figure C2: Pooled Datasets Agent Language and Customer Satisfaction



Red line: Affective Language; Blue line: Cognitive Language
Dotted lines: pointwise 95% confidence intervals

WEB APPENDIX D: EXPERIMENTS

Study 3 Stimuli

Below is the transcript from a real recording of a phone customer service conversation. We've changed the name of the company and redacted (removed) the customer's name if they said it.

Please read the conversation carefully, imagining that you were the customer involved.

Conver-sation	Topic	Speaker	Control	Treatment (Affective Words in Red Bold)
1	Return	Employee:	Shopsite customer service, my name is Ashley. How can I assist you?	**Thanks** for calling Shopsite customer service, my name is Ashley. I **hope** you're having a **nice** day. How can I **help** you?
1	Return	You:	Hi there. Hi. I'm trying to exchange an item that I got from Shopsite. And I'm trying to do it through the website, but I'm having some difficulty. It keeps telling me that my character max exceeded in the notes that I originally included and I can't seem to get past that or change that.	Hi there. Hi. I'm trying to return and exchange an item that I got from Shopsite. And I'm trying to do it through the website, but I'm having some difficulty. It keeps telling me that my character max exceeded in the notes that I originally included and I can't seem to get past that or change that.
1	Return	Employee:	Is this like for a gift message?	That's **frustrating**. Is this like for a gift message?
1	Return	You:	That's ok. So I bought a gift. I also bought a jacket for myself and that I need to swap out ummm because I actually think it's a little defective. And I'm trying to do that through the website and it's flagging me telling me that 'maximum exchange total of 250 was exceeded'. I'm not quite sure what's going on but I can't do it through the website.	That's ok. So I bought a gift. I also bought a jacket for myself and that I need to swap out ummm because I actually think it's a little defective. And I'm trying to do that through the website and it's flagging me telling me that 'maximum exchange total of 250 was exceeded'. I'm not quite sure what's going on but I can't do it through the website.
1	Return	Employee:	I can go ahead and process this for you today. May I have the order number that you're calling in regards to?	I'd be more than **happy** to go ahead and **help** with this for you today. May I **please** have the order number that you're calling in regards to?
1	Return	You:	Ah sure, let see, it's 4536901.	Ah sure, let see, it's 4536901.

1	Return	Employee:	And **may** I have your first and last name? [redacted]. Ok, thank you so much.	And may I **please** have your first and last name? [redacted]. Ok, that's **good**. Thank you so much.
1	Return	You:	Ahhh, the original shipping address for this order, let's see.	Ahhh, the original shipping address for this order, let's see.
1	Return	Employee:	Ok perfect. Thank you so much. So what item is it that you're trying to are you trying to return both or?	Ok perfect. Thank you so much. So what item is it that you're trying to are you trying to return both or?
1	Return	You:	No no just the jacket.	No no just the jacket.
1	Return	Employee:	Okay. And were you. I'm sorry?	Okay. And were you. I'm sorry?
1	Return	You:	Go ahead no I'm sorry	Go ahead no I'm sorry
1	Return	Employee:	Were you exchanging it for a different size or for a different item all together or?	Were you exchanging it for a different size or for a different item all together or?
1	Return	You:	Actually, we'd like another one of the same size.	Actually, we'd like another one of the same size.
1	Return	Employee:	Ok, alright. So let's go ahead and see. And may I ask the reason for the return today?	Ok, alright. So let's go ahead and see. And may I ask the reason for the return today?
1	Return	You:	Umm the tongue in the between the two shoes is a different length.	Umm the tongue in the between the two shoes is a different length.
1	Return	Employee:	Oh, oh no. I'm sorry about that.	Oh, oh no. I'm sorry about that.
1	Return	You:	Yeah, and they're a little expensive.	Yeah, and they're a little expensive.
1	Return	Employee:	Yeah. I can totally understand. I'd be more than happy to get those replaced for you.	Yeah. I can totally understand. I'd be more than happy to get those replaced for you.
1	Return	You:	No if you can send	No if you can send
1	Return	Employee:	I'm sorry, what was that address? Ok, perfect. So what will happen let me just confirm that email address for you.	I'm sorry, what was that address? Ok, perfect. So what will happen let me just confirm that email address for you.
1	Return	You:	Ummm, sure, yeah. And can I just umm how do I send them back. Do I just give them to the…..	Ummm, sure, yeah. And can I just umm how do I send them back. Do I just give them to the…..
1	Return	Employee:	No, I had. Alright perfect. So you'll receive the replacement pair as of tomorrow. And in regard to the pair that were defective, you have 14 days to send them back to us. Ummm, so let's go ahead and see. Now do you have access to a printer so I can email you a return label?	No, I had. Alright perfect. So you'll receive the replacement pair as of tomorrow. And in regard to the pair that were defective, you have 14 days to send them back to us. Ummm, so let's go ahead and see. Now do you have access to a printer so I can email you a return label?
1	Return	You:	Ok, ok perfect.	Ok, ok perfect.

1	Return	Employee:	We will issue you a prepaid return label and it will be UPS, so you will just drop them off at a UPS location.	We will issue you a prepaid return label and it will be UPS, so you will just drop them off at a UPS location.
1	Return	You:	That’s ok. I totally understand.	That’s ok. I totally understand.
1	Return	Employee:	Ok. One moment. Ok. Alright perfect. My system sometimes it works faster unless it’s |Inaudible|.	Ok. One moment. Ok. Alright perfect. My system sometimes it works faster unless it’s |Inaudible|.
1	Return	You:	That’s ok. I totally understand.	That’s ok. I totally understand.
1	Return	Employee:	Alright, perfect. So very shortly you’ll receive two emails from me, three emails. The first will be a return confirmation and the second will be an exchange and the third one will actually be the return label itself. And you’ll just click the link in that email to generate the return label and you’re be all sent.	Alright, perfect. So very shortly you’ll receive two emails from me, three emails. The first will be a return confirmation and the second will be an exchange and the third one will actually be the return label itself. And you’ll just click the link in that email to generate the return label and you’re be all sent.
1	Return	You:	Ok, great.	Ok, great.
1	Return	Employee:	And while I have you on the line, were there any other questions that I can assist you with today?	**Alright perfect.** And while I have you on the line, were there any other **problems** that I can **help** you with today?
1	Return	You:	No, no you’ve done great.	No, no you’ve done great.
1	Return	Employee:	Well we appreciate you shopping with Shopsite.	That's **excellent**. We’ll we appreciate you shopping with Shopsite. You have a wonderful day.
1	Return	You:	Thank you too	Thank you too
1	Return	Employee:	Buh-bye.	**Thank** you. You have a **great** day now!
1	Return	You:	Buh Bye.	Buh Bye.
2	Return	Employee:	Shopsite VIP. My name is Chris. What can I do for you today?	**Thank** you for calling Shopsite VIP. My name is Chris. How may I **help** you today?
2	Return	You:	I guess Sunday I called and ordered some shoes and then when I got the email about them they had screwed up the order and sent the wrong size. So, I called early this morning and they were going to stop the shipment. Ahhh, then further compounding the screw up, they went ahead and shipped them anyway, so I just refuse those or take it and then call you all to get	I guess Sunday I called and ordered some shoes and then when I got the email about them they had screwed up the order and sent the wrong size. So, I called early this morning and they were going to stop the shipment. Ahhh, then further compounding the screw up, they went ahead and shipped them anyway, so I just refuse those or take it and then call you all to get an authorization to send them back?

			an authorization to send them back?	
2	Return	Employee:	Ummm, let me see here. Do you have an order number?	Ummm, first let me say I **feel** your **pain** here. Do you have an order number so I can **help**?
2	Return	You:	Let's see here. Well, it's not on this... maybe it is here	Let's see here. Well, it's not on this... maybe it is here
2	Return	Employee:	That's, ummm. Did you request that the package was held?	That's **okay**. Did you request that the package was held?
2	Return	You:	Well, when I saw the thing, see I ordered a pair of size 14. Well, they sent UK 14, which is a US men's 15, which is too big for me. So, I called this morning and I had said this is all screwed up can you just stop the shipment since it hadn't shipped.	Well, when I saw the thing, see I ordered a pair of size 14. Well, they sent UK 14, which is a US men's 15, which is too big for me. So, I called this morning and I had said this is all screwed up can you just stop the shipment since it hadn't shipped.
2	Return	Employee:	That's right. So, just so you know that email goes out automatically when your order ships. To be honest, we cannot stop a shipment, but I can do..	That's **fair**. So, just so you know that email goes out automatically when your order ships. To be honest, we cannot stop a shipment, but I can **help.**
2	Return	You:	Well, you had plenty of time to stop it but nobody bothered to do it.	Well, you had plenty of time to stop it but nobody bothered to do it.
2	Return	Employee:	What I'm trying to say is we can't stop a shipment once somethings released to the fulfillment center. But there's a way around that. So like we're trying to return to sender, but it looks like you are holding the package.	What I'm trying to say is we can't stop a shipment once somethings released to the fulfillment center. But there's a way around that. So like we're trying to return to sender, but it looks like you are holding the package.
2	Return	You:	No. That package is not here.	No. That package is not here.
2	Return	Employee:	No, I understand what you're saying, but did you do anything through UPS?	No, I understand what you're saying, but did you do anything through UPS?
2	Return	You:	No.	No.
2	Return	Employee:	Okay.	Okay.

2	Return	You:	You just shipped it today. It couldn't be here already.	You just shipped it today. It couldn't be here already.
2	Return	Employee:	No, I understand that it's not going to be there, but ummm, but it will, I mean we did request it be returned to sender.	No, I understand that it's not going to be there, but ummm, but it will, I mean we did request it be returned to sender.
2	Return	You:	Ok. So if it shows up here, when it shows up here tomorrow, I just refuse it and let them take it back?	Ok. So if it shows up here, when it shows up here tomorrow, I just refuse it and let them take it back?
2	Return	Employee:	Well, to be honest it shouldn't show up there tomorrow because we did the return to sender.	Well, to be honest it shouldn't show up there tomorrow because we did the return to sender.
2	Return	You:	Ok. Well, it says it's going to so if it does do I just refuse it and send it back to you? I mean if it goes as smoothly as everything else has so far, it's probably going to come through the door here.	Ok. Well, it says it's going to so if it does do I just refuse it and send it back to you? I mean if it goes as smoothly as everything else has so far, it's probably going to come through the door here.
2	Return	Employee:	Ummm just a moment.	Ummm just a moment.
2	Return	You:	That's what I get for ordering over the damn internet.	That's what I get for ordering over the damn internet.
2	Return	Employee:	I'm very sorry about that sir, but you shouldn't you shouldn't be receiving your package.	I'm very sorry about that sir, but you shouldn't you shouldn't be receiving your package.
2	Return	You:	Ok. But if I,…. So far everything that's happened about it has been messed up. If it continues to be messed up and it comes through the door in my receiving dock, do I tell them to take it back or do I take it and then call you all to send it back.	Ok. But if I,…. So far everything that's happened about it has been messed up. If it continues to be messed up and it comes through the door in my receiving dock, do I tell them to take it back or do I take it and then call you all to send it back.
2	Return	Employee:	Ummm…. We could provide you with a shipping label to return it to us.	Ummm…. We could provide you with a shipping label to return it to us.
2	Return	You:	Yeah.	Yeah.
2	Return	Employee:	I think that it's going to be returned to us without it being received by you.	I think that it's going to be returned to us without it being received by you.
2	Return	You:	Okay.	Okay.
2	Return	Employee:	Umm but if but if you do receive it umm just just give us a buzz umm	Umm but if but if you do receive it umm just just give us a buzz umm
2	Return	You:	Go ahead and take it and fiddle around sending it back.	Go ahead and take it and fiddle around sending it back.

2	Return	Employee:	But it, I mean. I don't how UPS works if you can just be like I don't want the package return it to send.	But it, I mean. I don't how UPS works if you can just be like I don't want the package return it to send.
2	Return	You:	Yeah, I can refuse it.	Yeah, I can refuse it.
2	Return	Employee:	Ok. To be honest, yes you can do that if UPS accepts that, but I've never thought of that before but I'm sure it happens.	Ok. To be honest, yes you can do that if UPS accepts that, but I've never thought of that before but I'm sure it happens.
2	Return	You:	Ok, that's what I'm counting on then.	Ok, that's what I'm counting on then.
2	Return	Employee:	Correct.	I **hope** it works.
2	Return	You:	Alright, thank you.	Alright, thank you.
2	Return	Employee:	It's nothing.	My **pleasure**.
2	Return	You:	Bye	Bye
3	Order	Employee:	Thank you for calling Shopsite VIP. This is Ali speaking, how can I assist you?	Thank you for calling Shopsite VIP. This is Ali speaking, how **may** I **help** you?
3	Order	You:	Hi, um, I was trying to order something, uh, the day before yesterday. And I tried to change the expiration date on my Visa.	Hi, um, I was trying to order something, uh, the day before yesterday. And I tried to change the expiration date on my Visa.
3	Order	Employee:	Mm-hmm.	**Okay.**
3	Order	You:	And payment information because I got a new Visa. And now I've got it all screwed up. It's not-it says I need help.	And payment information because I got a new Visa. And now I've got it all screwed up. It's not-it says I need help.
3	Order	Employee:	Um, what is your email address?	Okay, I'd **love** to **help**. What is your email address?
3	Order	You:	Yeah.	Yeah.
3	Order	Employee:	Okay. Got your account here. And then for security purposes, can I get you to verify your shipping address? And the city and zip code?	Okay, **perfect**. Oh, **great**. Got your account here. And then for security purposes, can I get you to verify your shipping address? And the city and zip code?
3	Order	You:	And what?	And what?
3	Order	Employee:	|Redacted|	|Redacted|
3	Order	You:	|Redacted|	|Redacted|
3	Order	Employee:	Oh okay. So let's go ahead.	Oh okay. So let's go ahead.
3	Order	You:	I tried to, um, delete some of the other cards that were on there. I don't know.	I tried to, um, delete some of the other cards that were on there. I don't know.

3	Order	Employee:	That's okay. So let's go ahead and try to edit this information this card. Um, let's see. And how does your name appear on the card? Perfect.	That's okay. So let's go ahead and try to edit this information this card. Um, let's see. And how does your name appear on the card? Perfect.
3	Order	You:	Uh, yeah.	Uh, yeah.
3	Order	Employee:	Alright. So I do need to recollect the card number when we-when we try to, um, adjust any information on the payment information. So	Alright. So I do need to recollect the card number when we-when we try to, um, adjust any information on the payment information. So
3	Order	You:	You want me to give you the card number?	You want me to give you the card number?
3	Order	Employee:	So the way that we do it is actually kind of unique in our system. We don't like to have the card number said over the phone with us because	So the way that we do it is actually kind of unique in our system. We don't like to have the card number said over the phone with us because
3	Order	You:	Right.	Right.
3	Order	Employee:	Yeah, the calls are recorded. So what, uh, we do is we send you over to the card line and you just type in the card number on the keypad of the phone that you're using to talk to me. And then	Yeah, the calls are recorded. So what, uh, we do is we send you over to the card line and you just type in the card number on the keypad of the phone that you're using to talk to me. And then
3	Order	You:	Okay, okay.	Okay, okay.
3	Order	Employee:	Yeah, once you're done, you'll hit the # sign and it'll send you and the card number back over to me, okay?	Yeah, once you're done, you'll hit the # sign and it'll send you and the card number back over to me, okay?
3	Order	You:	Okay.	Okay.
3	Order	Employee:	Cool. So let me send you on over.	Cool. So let me send you on over.
3	Order	You:	Okay	Okay
3	Order	Employee:	Alright, perfect. I do see those last four digits so that's over here. So let me go ahead and save that for you.	Alright, perfect. I do see those last four digits so that's over here. So let me go ahead and save that for you.
3	Order	You:	Okay, great. Thank you.	Okay, great. Thank you.
3	Order	Employee:	Yeah.	Yeah.
3	Order	You:	So then it'll change ultimately on the payment information?	So then it'll change ultimately on the payment information?
3	Order	Employee:	Yes, so when you go to use it, it show should everything as	Yes, so when you go to use it, it show should everything as
3	Order	You:	Okay, right now it doesn't. But	Okay, right now it doesn't. But

3	Order	Employee:	Okay, so, um, if you click on my account and then maybe go back	Okay, so, um, if you click on my account and then maybe go back
3	Order	You:	Oh okay. Let me see. Yes, it does on that. It just doesn't-I'm trying-I was trying to order something and then my order didn’t go through because it said there was a problem with my credit card. And I knew what it was because I was messing around trying to edit it |Laugh|.	Oh okay. Let me see. Yes, it does on that. It just doesn't-I'm trying-I was trying to order something and then my order didn’t go through because it said there was a problem with my credit card. And I knew what it was because I was messing around trying to edit it |Laugh|.
3	Order	Employee:	Right. Yeah, that's okay. It happens sometimes. Um, okay.	Right |Laugh|. Yeah, that's okay. It happens sometimes. Um, okay.
3	Order	You:	Okay.	Okay.
3	Order	Employee:	So now when you go to check out, it should work.	So now when you go to check out, it should work.
3	Order	You:	Okay, alright. Let's see. Let me see my card. Proceed to checkout. Yup, thank you.	Okay, alright. Let's see. Let me see my card. Proceed to checkout. Yup, thank you.
3	Order	Employee:	Awesome. Sweet.	Awesome. Sweet.
3	Order	You:	Thank you.	Thank you.
3	Order	Employee:	Yeah, no problem.	Yeah, **please**, no **worries** at all.
3	Order	You:	Thanks.	Thanks.
3	Order	Employee:	Anytime.	My **pleasure**.
3	Order	You:	Okay. Thank you very much.	Okay. Thank you very much.
3	Order	Employee:	You too.	**Thank** you too.
3	Order	You:	Bye-bye. Bye-bye.	Bye-bye. Bye-bye.
3	Order	Employee:	Bye.	Bye. Have a **nice** day.
4	Shipping	Employee:	This is Alisha at Shopsite. How are you doing today?	**Good** day. This is Alisha at Shopsite. **Thank** you for calling. How are you doing today?
4	Shipping	You:	I'm okay.	I'm okay.
4	Shipping	Employee:	What can I do for you?	What can I **help** you with?
4	Shipping	You:	Um, I ordered a pair of slippers on Thursday night, and it said - I chose one-day shipping. And it's not going to be here until the 8th now, so I'm a little frustrated because it was supposed to be for a birthday present today.	Um, I ordered a pair of slippers on Thursday night, and it said - I chose one-day shipping. And it's not going to be here until the 8th now, so I'm a little frustrated because it was supposed to be for a birthday present today.
4	Shipping	Employee:	Let me take a look. Do you happen to have that order number?	**Okay, thanks.** I can **help** with that. Do you happen to have that order number?
4	Shipping	You:	Um, yes. Hold on. It is 7854359.	Um, yes. Hold on. It is 7854359.
4	Shipping	Employee:	Okay, let me pull that up. Give me one second.	Okay, let me pull that up. Give me one second.
4	Shipping	You:	I never got another email about it either, like you said you would.	I never got another email about it either, like you said you would.

4	Shipping	Employee:	Okay, so let me see, can you just verify your shipping address for me? /	Okay, so let me see, can you just verify your shipping address for me? /
4	Shipping	You:	Yes.	Yes.
4	Shipping	Employee:	Okay. So they were ordered Thursday after our cut-off time, which means the first day they'd process was Friday. However, the reason you're getting them on Monday is because UPS doesn't ship on Saturday or Sunday.	Okay. So they were ordered Thursday after our cut-off time, which means the first day they'd process was Friday. However, the reason you're getting them on Monday is because UPS doesn't ship on Saturday or Sunday.
4	Shipping	You:	Well -	Well -
4	Shipping	Employee:	- no delivery, so the next business day is –	- no delivery, so the next business day is -
4	Shipping	You:	I get - I get UPS orders on Saturday.	I get - I get UPS orders on Saturday.
4	Shipping	Employee:	We're not - UPS isn't contracted with us to ship on Saturday.	We're not - UPS isn't contracted with us to ship on Saturday.
4	Shipping	You:	So it's very frustrating that wasn't explained when I, when I purchased my order, and I didn't get a follow-up email telling me that otherwise I would have cancelled the order and gone through Amazon.	So it's very frustrating that wasn't explained when I, when I purchased my order, and I didn't get a follow-up email telling me that otherwise I would have cancelled the order and gone through Amazon.
4	Shipping	Employee:	Um, the next - it's next business day shipping, so our cut-off time is 1 PM, that's why it's pushed to Monday, because it got processed on Friday, and then the next business day would be Monday.	Yes, I'm **sorry** about that. Um, the next - it's next business day shipping, so our cut-off time is 1 PM, that's why it's pushed to Monday, because it got processed on Friday, and then the next business day would be Monday.
4	Shipping	You:	Yeah. Okay. Thank you.	Yeah. Okay. Thank you.
4	Shipping	Employee:	Have a great day.	**Thanks** again for your **patience**. I **hope** you'll shop with us again.
4	Shipping	You:	Bye.	Bye.
4	Shipping	Employee:	Bye, bye.	**Okay.** Have a **nice** day.
5	Product	Employee:	You're calling Shopsite. My name is Chuck. How can I help you today?	**Thank** you for calling Shopsite. My name is Chuck. How can I help you today?
5	Product	You:	Hi, good morning, Chuck. Can you tell me if a particular shoe is in stock. I got an email at 3.20 this morning, but it looks like it might already be gone. This -	Hi, good morning, Chuck. Can you tell me if a particular item is in stock. I got an email at 3.20 this morning, but it looks like it might already be gone. This –

5	Product	Employee:	Let me see. Yeah, what's the SKU number?	**Happy** to **help**. **Okay**, what's the SKU number?
5	Product	You:	783442923	783442923
5	Product	Employee:	The Bailey mini skirt. Looks like.	The Bailey mini skirt. Looks **nice**.
5	Product	You:	Yes. In a black and size 7.	Yes. In a black and size 7.
5	Product	Employee:	Looks like we do have it in stock.	**Excellent**. it looks like we do have it in stock.
5	Product	You:	OK.	OK.
5	Product	Employee:	You're going to place that order?	You're going to place that order?
5	Product	You:	Yes, please.	Yes, please.
5	Product	Employee:	Alright, what's your email address?	Alright, what's your email address?
5	Product	You:	|Redacted|	|Redacted|
5	Product	Employee:	Great. Let me pull this up. And who am I speaking with today?	Great. Let me pull this up. And who am I speaking with today?
5	Product	You:	|Redacted|	|Redacted|
5	Product	Employee:	And what city, state and zip code?	And what city, state and zip code?
5	Product	You:	|Redacted|	|Redacted|
5	Product	Employee:	Let's put this order in here.	Let's put this order in here.
5	Product	You:	It said only one left. I liked it.	It said only one left. I liked it.
5	Product	Employee:	Mm-hmm.	Mm-hmm.
5	Product	You:	Gone, so . . . I didn’t know what that meant.	Gone, so . . . I didn’t know what that meant.
5	Product	Employee:	Is it gone?	Is it gone?
5	Product	You:	Uh-huh. Like it, click it here to buy.	Uh-huh. Like it, click it here to buy.
5	Product	Employee:	How come it said - how come that's saying it's gone. Now it said that size 7 is gone. Let me see something. Why did it say it was available a second ago, let me see. Let me check here. Let me see. There's black. Uh, now it says we don’t have it. Let me see here.	How come it said - how come that's saying it's gone. Now it said that size 7 is gone. Let me see something. Why did it say it was available a second ago, let me see. Let me check here. Let me see. There's black. Uh, now it says we don’t have it. Let me see here.
5	Product	You:	OK. That's what I did. I looked at it and it was gone before my very eyes. I thought "What is this?"	OK. That's what I did. I looked at it and it was gone before my very eyes. I thought "What is this?"
5	Product	Employee:	Yeah, it's so weird.	Yeah, it's so weird.
5	Product	You:	It did that last week.	It did that last week.
5	Product	Employee:	I wonder if someone just clicked the order now. Yeah. Do you want to put a notify email for you for when it comes back.	I wonder if someone just clicked the order now. Yeah. Do you want to put a notify email for you for when it comes back.
5	Product	You:	Yes.	Yes.
5	Product	Employee:	Alright, let's do that.	Alright, let's do that.

5	Product	You:	Third time's the charm.	Third time's the charm.
5	Product	Employee:	Yeah, let's do that. I know these skirst and this time of year, they're like really popular and these things sell so fast I'll put you on the notify list again so hopefully this time it'll work out, um.	Yeah, let's do that. I know these skirst and this time of year, they're like really popular and these things sell so fast I'll put you on the notify list again so hopefully this time it'll work out, um.
5	Product	You:	Well, the time of notification is the same. It's 3:20 pm as it was last week.	Well, the time of notification is the same. It's 3:20 pm as it was last week.
5	Product	Employee:	Mm-hmm.	Mm-hmm.
5	Product	You:	But maybe somehow that message just got rolled over into something -	But maybe somehow that message just got rolled over into something -
5	Product	Employee:	Really. I put you on the same - the notify list again, but wait, maybe I can do something more. You know, um, also it looks like Amazon has it in stock.	That's not **great**. OK I put you on the same - the notify list again, but wait, maybe I can **help** even more. You know, um, also it looks like Amazon has it in stock.
5	Product	You:	Oh, I'd rather just deal with y'all.	Oh, I'd rather just deal with y'all.
5	Product	Employee:	That's fine.	That's **nice** of you.
5	Product	You:	Just easier that way.	Just easier that way.
5	Product	Employee:	So yeah, you'll get the email again once it comes in stock.	**Excellent,** whatever's **best** for you**.** So you'll get the email again once it comes in stock.
5	Product	You:	OK, thank you.	OK, thank you.
5	Product	Employee:	You bet. Bye bye.	My **pleasure**. Have a **nice** day.
5	Product	You:	Bye.	Bye.

(Words from the affective process dictionary are presented in ***red*** *bold font.)*

Study 3 Manipulated Affective Language and Resultant Satisfaction Means by Stimuli Version

Stimuli Version	Call Topic	N	Affective Language (LIWC) Control	Affective Language (LIWC) Treatment	Affective Language (LIWC) Difference	Customer Satisfaction Means Control	Customer Satisfaction Means Treatment
1	Return	134	0.88	4.47	3.60	6.09	6.48
2	Return	134	0.36	2.76	2.40	3.16	3.62
3	Order	143	0.35	4.64	4.29	3.74	4.11
4	Shipping	132	2.67	6.25	3.58	5.72	6.03
5	Product	143	1.16	5.05	3.90	4.47	5.28

Study 3 Results without the Exclusion

Exclusion. Study 3 and all subsequent experiments used the words per minute exclusion approach described in this study's pre-registration (https://aspredicted.org/M1K_4VC). Specifically, we exclude participants who move through the main stimuli page (the conversation transcript and dependent variable measure) at a time interval consistent with 500 words per minute (WPM) or greater (based on the number of words on the main stimuli page) according to a Qualtrics timer that is not observable to participants. The 500 WPM exclusion rule has been used by the first author in all laboratory studies conducted over the last 13 years. It is based on published guidelines on average adult reading speed and comprehension. Normal adult reading rates for comprehension are 200-250 WPM. Five hundred WPM captures more than three standard deviations (>99.7%) of adult readers (Just and Carpenter 1987).

We preregistered a targeted final sample size of 75 participants per condition after the exclusion, but mistakenly pre-registered the design as a 2 (control vs. warm start and end treatment) x 5 (stimuli sampling) for a total 10 conditions, resulting in a total target N of 750, and proceeded to collect on that basis. Of course, there was actually only one control condition for comparison against the 5 treatment condition stimuli (not five of the same control conditions), so there were actually only six conditions.

To account for exclusions, we asked Prolific for 788 participants (750 + 5%). Seven hundred and forty eight participants actually completed the study. After the standard exclusion, 686 participants remained, corresponding to an exclusion rate of 8%. Achieved power after exclusion was 86% ($\alpha = .05$). Results following the pre-registered procedure are reported in the main manuscript. As shown below, all results replicate without the exclusion.

Results without the Exclusion. As predicted, across a range of real customer service conversations from Study 1, analysis of variance reveals that using more affective language at the start and end enhances customer satisfaction (Treatment = 5.11 vs. Control = 4.63; $F(1, 746) = 12.95$, $p < .001$, $\eta^2_p = .02$).

Results remain the same controlling for potential differences across the five conversations and the interaction of condition with conversation ($F(1, 738) = 19.99$, $p < .001$, $\eta^2_p = .03$). As for the controls, there was an irrelevant main effect of conversation on satisfaction ($F(4, 738) = 92.76$, $p < .001$). More importantly, the beneficial effect of adding more affective language to the start and end did not vary significantly across the different conversations (condition x conversation interactoin $F(4, 738) = .53$, $p = .715$).

Study 4A Experimental Stimuli

Imagine you called customer service at Shopsite, an online retailer, and this was the conversation you had with a service agent:

Agent:	Hello. [How **might** I **assist** you today? / I **hope** you're **enjoying** this fine day?]
You:	I can't figure out how to get the free shipping.
Agent:	I think I can find a solution. I know it can be a little complex to locate. I'll explain where… scroll down a bit. See the dropdown menu at the bottom right?
You:	Uh… ok. I got it.

Agent: [I **guess** your issue is **resolved** then? / I **hope** your issue is **OK** then?]

You: Yes, thank you. Bye now.

(Words from the affective process and cognitive process dictionaries are presented in ***red*** *and* ***blue*** *bold font, respectively.)*

Study 4A LIWC Values for Manipulated Agent Turns

Condition	Turn Language	LIWC Dictionary Measures: Affective	LIWC Dictionary Measures: Cognitive	Net Language (+ = More Affective, - = More Cognitive)
Cognitive only	Hello. How might I assist you today?	0.00	28.57	-28.57
Dynamic	Hello. I hope you're enjoying this fine day?	37.50	12.50	25.00
Cognitive only	I guess your issue is resolved then?	14.29	28.57	-14.28
Dynamic	I hope your issue is OK then?	28.57	14.29	14.28

Study 4A Results without the Exclusion

Exclusion. Study 4A aimed to have a final sample size of 75 participants per condition after the standard exclusion for a total target N = 150. To account for the standard exclusion, we asked Amazon Mechanical Turk for 173 participants (150 + 15%). One hundred sixty eight participants actually completed the study. After the standard exclusion, 146 participants remained, corresponding to an exclusion rate of 13%. Achieved power after exclusion was 89% (α = .05). Results following the standard exclusion are reported in the main manuscript. As shown below, all results replicate without the exclusion.

Results without Exclusion. As predicted, manipulating agent language based on our dynamic recommendation (i.e., more affective language at the start and end) improved customer satisfaction ($M_{dynamic}$ = 6.19, $SD_{dynamic}$ = .95 vs. $M_{all\ cognitive}$ = 5.82, $SD_{all\ cognitive}$ = .87; $F(1, 166)$ = 6.92, p = .009, η^2_p = .04). It also led agents to be perceived as marginally more helpful ($M_{dynamic}$ = 6.10, $SD_{dynamic}$ = .85 vs. $M_{all\ cognitive}$ = 5.84, $SD_{all\ cognitive}$ = .90; $F(1, 166)$ = 3.45, p = .065, η^2_p = .02).

Study 4B Stimuli

Imagine you called customer service at Shopsite, an online retailer, and this was the conversation you had with a service agent:

Agent: Hello. [How **might** I **assist** you today? / I **hope** you're **enjoying** this fine day?]

You: I can't figure out how to get the free shipping.

Agent: [I **think** I can **find** a solution. I **know** it can be a little **complex** to find. / I **appreciate** how **annoying** that can be. I've been frustrated locating it myself.]

[I'm **glad** you called… / I'll **explain** where…]

…just scrolling down a bit….

[You **should find** that button at the bottom right. / You'll feel better with that button at the bottom right.]

You: Uh… ok. I got it.

Agent: [I **guess** your issue is **resolved** then? / I **hope** your issue is **OK** then?]

You: Yes, thank you. Bye now.

*(Words from the affective process and cognitive process dictionaries are presented in **red** and **blue** bold font, respectively.)*

Study 4B LIWC Values for Additional Manipulated Agent Language beyond Study 4A (i.e., manipulation of Middle 50% to support the total of eight experimental conditions)

		LIWC Dictionary Measures		Net Language (+ = More Affective, - = More Cognitive)
Condition	Middle 50% Language	Affective	Cognitive	
Cognitive	I think I can find a solution. I know it can be a little complex to find. I'll explain where. You should find that button at the bottom right.	0.00	31.03	-31.03
Affective	I appreciate how annoying that can be. I've been frustrated locating it myself. I'm glad you called. You'll feel better with that dropdown menu at the bottom right.	17.86	10.71	7.15

Study 4B Results without the Exclusion

Exclusion. Study 4B aimed to have a final sample size of 75 participants per condition after the standard exclusion for a total target N = 600. To account for the exclusion, we asked Prolific for 630 participants (600 + 5%). Six hundred thirty one participants actually completed the study. After the standard exclusion, 603 participants remained, corresponding to an exclusion rate of 4%. Achieved power after exclusion was 85% or greater across the seven key contrasts (α = .05). Results following the standard exclusion are reported in the main manuscript. As shown below, all results replicate without the exclusion.

Results without Exclusion. Results indicate that language based on the dynamic model's recommendation improved customer satisfaction (M = 5.25, SD = 1.66) relative to all other conditions. This includes the competence throughout recommendation of prior research (M = 4.58, SD = 2.00; $F(1, 156) = 7.20$, $p = .008$, $\eta^2_p = .05$) as well as warm only at the start (M =

4.663, SD = 2.01; $F(1, 156) = 4.13$, $p = .044$, $\eta^2_p = .03$), warm only at the end (M = 4.11, SD = 2.00; $F(1, 156) = 15.18$, $p < .001$, $\eta^2_p = .09$), warm throughout (M = 4.04, SD = 1.93; $F(1, 156) = 17.99$, $p < .001$, $\eta^2_p = .10$), competence-warmth-competence (M = 4.27, SD = 2.08; $F(1, 156) = 10.88$, $p < .001$, $\eta^2_p = .07$), competence-competence-warmth (M = 4.11, SD = 2.00; $F(1, 156) = 15.18$, $p < .001$, $\eta^2_p = .09$), and warmth-competence-competence (M = 4.66, SD = 2.01; $F(1, 156) = 4.13$, $p = .044$, $\eta^2_p = .03$).[5] These findings underscore the notion that the specific dynamic sequence from our theorizing is superior to a variety of alternative sequences, and further supports prior empirical support that *when* language is used matters (rather than merely *what* language is used).

Study 5 Stimuli

Speaker	Control (competent-warm-competent)	Treatment (warm-competent-warm)
Agent:	This is JetAir service. How **might** I **assist** you?	**Thanks** for calling JetAir service. How can I **help** you?
You:	Why was my flight cancelled?	Why was my flight cancelled?
Agent:	Oh, I can **answer** that. It was…	**Ugh**, I'm **sorry** about that. It was..
You:	Just get me on a new flight by 3pm. My booking reference is AE3XH.	Just get me on a new flight by 3pm. My booking reference is AE3XH.

[5] As in Study 4, we also replicate the results using the Study 1 retailer's satisfaction measure "How helpful was the agent?". Our dynamic treatment condition again outperformed the recommendation of prior research ($M_{dynamic} = 5.54$, $SD_{dynamic} = 1.58$ vs. $M_{all\ cognitive} = 4.87$, $SD_{all\ cognitive} = 2.02$; $F(1, 146) = 5.07$, $p = .026$, $\eta^2_p = .03$) and all six other conditions (all *p*s < .02; all $\eta^2_p > .03$).

Agent:	I'm **sorry**. One moment... I'm just **hoping** to **share** something that might be **alright** for you... **Thank** you for waiting… So… **luckily** I feel like I've found a **good** one. I can **gladly help** you get you a spot on the 3:15pm flight.	I **understand.** One moment... I'm just **trying** to **find** something that might **work** for you... Just a sec… So… I **think** I've **found** an **option**. I can **actually** get you a spot on the 3:15pm flight.
You:	If that's the best we can do.	If that's the best we can do.
Agent:	The weather delays at your destination have really **complicated everything**. I **acknowledge** this is not quite right for your needs, but I've tried to **find** a **solution**.	**Thank** you for your patience. The weather delay at your destination has been a real **nightmare**. I **appreciate** this is not quite right for your needs, but I've tried my **best** to **help.**
You:	Go ahead. Please give me the new flight info…	Go ahead. Please give me the new flight info…

(Words from the affective process and cognitive process dictionaries are presented in ***red*** *and* ***blue*** *bold font, respectively.)*

Study 5 LIWC Values for Manipulated Agent Turns and Full Interaction Level

		LIWC Dictionary Measures		Net Language (+ = More Affective, - = More Cognitive)
Condition	Turn Language	Affective	Cognitive	
Control	First 25%	0.00	18.75	-18.85
	Middle 50%	20.93	0	20.93
	Last 25%	0	18.50	-18.50
	Full Conversation	**9.0**	**8.0**	
Treatment	First 25%	23.53	0	23.53
	Middle 50%	0	22.22	-22.22
	Last 25%	15.15	0	15.15
	Full Conversation	**9.0**	**8.0**	

Study 5 Results without the Exclusion

As in prior studies and per the preregistration, we asked Prolific for 158 participants (150 + 5%) to account for exclusions. One hundred and fifty eight participants actually completed the study. After the standard exclusion, 154 participants remained, corresponding to an exclusion rate of 3%. Achieved power after exclusion was 86% ($\alpha = .05$). Results following the pre-registered procedure are reported in the main manuscript. As shown below, all results replicate without the exclusion.

Results

As predicted, agent language based on our dynamic recommendation (i.e., warmth-competence-warmth) improved customer satisfaction versus a fully reversed control (i.e., competence-warmth-competence; $M_{dynamic} = 5.77$, $SD_{dynamic} = 1.25$ vs. $M_{fully\ reversed} = 5.06$, $SD_{fully\ reversed} = 1.37$; $F(1, 156) = 11.53$, $p = .002$, $\eta^2_p = .07$).

WEB APPENDIX E: MODERATING ROLE OF SEVERITY

While the results of the studies reported in the main paper support our theorizing about how conversational time moderates the effect of affective and cognitive language, one could wonder whether other factors might further shape these relationships. Given all the different potential secondary moderators, and the many conversational moments over which one could explore them, doing so fully is beyond the scope of this paper. That said, to illustrate how one might approach such opportunities, we explore whether issue severity moderates the benefit of affective or cognitive language at particular conversational moments.

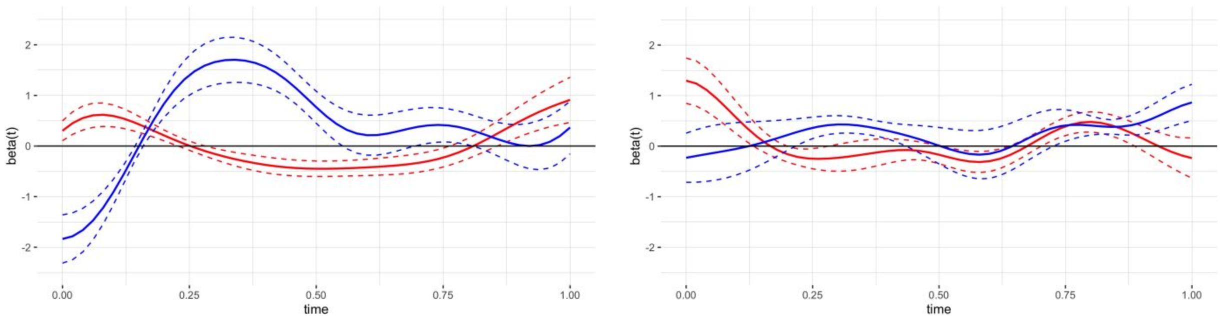
Study 1 Data

To begin to consider this possibility, we first examined the Study 1 field data. Judges had rated the severity of each call, so we use median severity to split the data and ran the full model (Model 3) for each half.

Severity appears to significantly moderate the sign and significance of cognitive language across different conversational moments. While cognitive language at the start (first 25%) is significantly more costly in higher severity interactions (negative significant beta coefficient) than lower severity interactions (null beta coefficient, compare blue lines in Figure E1 Panel A vs. B), for example, it is more beneficial at the end (last 25%) for lower severity interactions. Cognitive language remains beneficial in the middle of the interaction regardless of severity.

Figure E1: Beta Curves for Agent Affective and Cognitive Language
by Conversational Moment, Moderated by Issue Severity (Model 3)

Panel A: Higher Severity Panel B: Lower Severity



Red lines: Affective Language; Blue lines: Cognitive Language
Dotted lines represent pointwise 95% confidence intervals

Severity does not seem to play as much of a moderating role when it comes to affective language. Results indicate that affective language is beneficial (positive significant beta coefficient) for both higher and lower severity issues at the interaction's start and end (compare red lines in Figure E1 Panel A vs. B), with some differences in when during the end (last 25%) affective language is beneficial. As with our main results, affective language is not beneficial in the interaction's middle period regardless of severity.

Experimentally Testing Moderation by Severity

To more directly test severity's moderating role, we use an experiment. The field data most clearly suggest that (1) more cognitive language at the start should be detrimental for higher severity issues and (2) more cognitive language at the end should be beneficial for lower severity issues. Study 6 manipulates cognitive language at both of these moments and examines its causal impact.

Method. Participants (N = 806, Prolific) were randomly assigned to condition in a 2 (severity: higher vs. lower) x 3 (language: control vs. more cognitive at the start vs. more cognitive at the end) between-subjects designs. To further test robustness, we took a call from the airline field data (i.e., a customer asking for a refund for a cancelled flight), and manipulated agent language (see Web Appendix D for full stimuli for all conditions). This study was preregistered (https://aspredicted.org/CQP_GPQ).

First, we manipulated issue severity. In the high severity condition, participants were told "You're extremely frustrated and concerned because you have important plans today, and if you don't get on a plane in the next couple hours, you'll miss them." In the lower severity baseline condition, however, participants were told "You're not particularly frustrated or concerned because you don't have any plans today." This manipulation was used because we observed considerable heterogeneity in how frustrated customers appeared to be with a travel delay or cancellation in the field data based on whether they had an "important" or "urgent" reason for their travel. The manipulation worked as intended: participants in the high severity condition found the situation more severe ($M_{\text{higher severity}} = 4.66$, SD = 1.41 vs. $M_{\text{lower severity}} = 3.18$, SD = 1.53; $F(1, 804) = 206.48$, $p < .001$, $\eta^2_p = .21$).

Second, we manipulated agent language. In the control condition, language was similar to the original call. For the more cognitive at start condition, agent cognitive language was increased in the first 25% of the conversation. For the more cognitive at end condition, agent cognitive language was increased in the last 25% of the conversation. See Web Appendix D for complete stimuli and moment-to-moment affective and cognitive language values by condition.

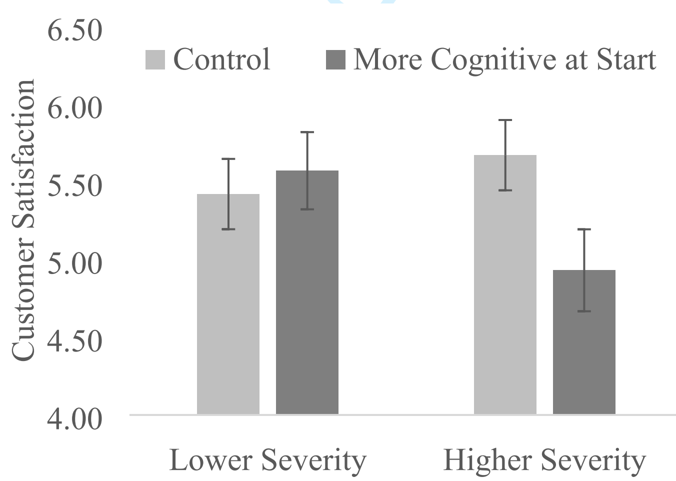
Finally, participants indicated their satisfaction using the Study 3-5 customer satisfaction dependent measure.

Main Results. Main effects of severity ($F(1, 800) = 9.92$, $p = .002$) and language ($F(1, 800) = 12.57$, $p < .001$), were qualified by the predicted severity x language interaction ($F(2,

800) = 8.44, $p < .001$). Following the pre-registration, we decomposed this interaction into the two focal subsets of the data for which we had predictions (i.e., more cognitive language at start versus control, and more cognitive language at end versus control).

Cognitive Language at Start. An effect of language ($F(1, 528) = 5.90$, $p = .015$) was qualified by the predicted severity x language interaction ($F(1, 528) = 12.87$, $p < .001$). Consistent with the moderation in the field data, while using more cognitive language at the start decreased customer satisfaction (relative to the control) when the issue was higher severity (M = 4.93, SD = 1.55 vs. M = 5.67, SD = 1.33, $F(1, 528) = 18.10$, $p < .001$, $\eta^2_p = .06$) it had no effect when the issue was lower severity (M = 5.57, SD = 1.46 vs. M = 5.42, SD = 1.33; $F(1, 528) = .67$, $p = .413$, $\eta^2_p = .00$; Figure E2).

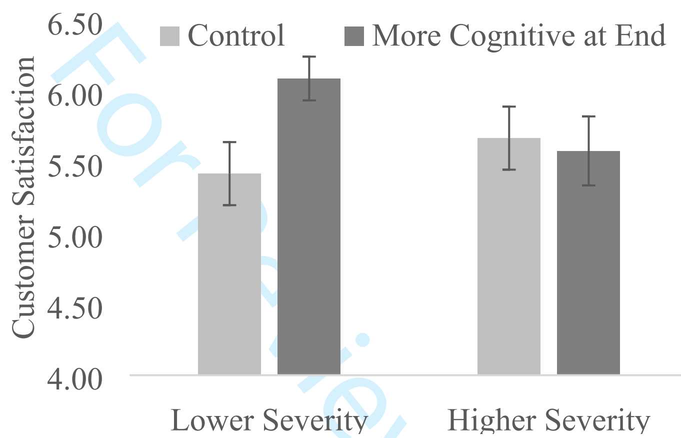
Figure E2: Impact of Cognitive Language at Start



Note: Error bars represent 95% confidence intervals

Cognitive Language at End. An effect of language ($F(1, 548) = 7.22$, $p = .007$) was qualified by the predicted severity x language interaction ($F(1, 548) = 12.02$, $p < .001$). Consistent with the moderation in the field data, while using more cognitive language at the end increased customer satisfaction (relative to the control) when the issue was lower severity (M = 6.09, SD = .92 vs. M = 5.42, SD = 1.33, $F(1, 548) = 19.44$, $p < .001$, $\eta^2_p = .08$), it had no effect when the issue was higher severity (M = 5.58, SD = 1.46 vs. M = 5.67, SD = 1.33; $F(1, 548) = .30$, $p = .584$, $\eta^2_p = .00$; Figure E3).

Figure E3: Impact of Cognitive Language at End



Note: Error bars represent 95% confidence intervals

Discussion. Results of Study 6 provide further evidence for the findings observed in the field. More cognitive language at the start was detrimental when the issue was more severe, and more cognitive language at the end was beneficial when the issue was less severe. Directly manipulating language at different conversational moments underscores its causal impact on customer satisfaction. Further, the results demonstrate that factors like issue severity can moderate how much different language features are beneficial at different points in a conversation.

As with any study, this one is not without limitations. Because only the Study 1 retailer field data set included a control variable that might be of potential interest as a moderator of the temporally moderated effect of affective and cognitive language on customer satisfaction, the field data analysis that opens this preliminary experimental investigation used the data from Study 1. Given there are only minor differences in the start-of-conversation functional forms across the retailer and airline data sets (cf. main manuscript Figures 2 and 4), and because we had not yet used an airline setting in experiments at the time we conducted this study, we thought it would be ideal use an airline service scenario. In pretests, we also found that we could more stably and significantly manipulate perceived severity with experimental participants using an airline than retailer service issue. That said, future work might seek to replicate the results of this study using retail service stimuli to address this shortcoming.

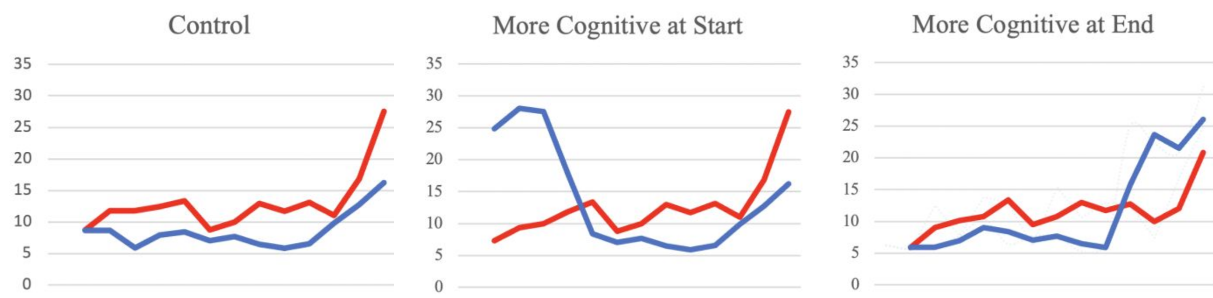
Study 6 Experimental Stimuli Development Detail

As described in the main study reporting above, for Study 6 we manipulated the agent language in a real airline service conversation. For the control condition, we sought to keep the moment-to-moment trend in use of cognitive and affective language similar to the mean of the field data (Figure C1). We then increased the agent's cognitive language in either the first 25% (more cognitive at start condition) or last 25% (more cognitive at end condition) while trying to maintain naturalism and minimize changes in overall language use and meaning.

Figure E4 presents the resultant moment-to-moment dynamic of affective and cognitive language for each experimental condition, followed by complete stimuli for each condition with affective and cognitive dictionary words highlighted in **red** and **blue**, respectively. One word (understand) appears in both dictionaries, and is indicated using both colors (i.e., **understand**).

Figure E4: Study 6 LIWC Values for Affective and Cognitive

Language (Y axis) over Stimuli Conversational Time (X axis)



Red line: Affective Language; Blue line: Cognitive Language

Study 6 Stimuli

Imagine you're in a car on the way to the airport. You call your airline's service number because your flight was just cancelled.

Lower severity condition: **You're not particularly frustrated or concerned because you don't have any plans today.**

Higher severity condition: **You're extremely frustrated and concerned because you have important plans today and if you don't get on a plane quickly you'll miss them.**

Here's the conversation you have with the airline's service agent:

Condition	Speaker	Conversational Language
Control	Agent:	This is Jet Airline customer **support**. My name is Charlie. **How** can I assist you today?
Control	You:	Hi, I was on a flight that was supposed to get out from here to Chicago this morning at 10:50. And I'm not thrilled it was cancelled.
Control	Agent:	That's **not** **good**. If you have a confirmation number, I'll do my **best** to **figure** out the issue.
Control	You:	Um, sure. But I don't think you're hearing me…
Control	Agent:	Yes, I hear you. Your flight time **changed**. I can **help** you with that for **sure**.
Control	You:	The confirmation number is J2Y5FZ, but I need to get on a new flight now. I need you to address this. I don't need an explanation or information.
Control	Agent:	Uh, yes I **appreciate** that. **Please** let me address this issue. I'll **need** the name on the ticket.

Control	You:	It's [your name].
Control	Agent:	**Alright**. One moment **please**. I'm just **trying** to **hopefully** **find** something **good** for you. OK, I've **found** it. Our system booked you on the next available flight. That's 3:15pm.
Control	You:	Hilarious. Um, that's not gonna work for me, so I'm actually on the way to the airport now to get a different flight, but I'm hoping you can credit me the amount I spent on this flight.
Control	Agent:	Let me see if I can take **care** of you... **Okay**, **because** your flight was **regrettably** delayed due to **bad** weather, I can **make** this into a flight credit for you.
Control	You:	No, I'd rather have a full credit. A refund.
Control	Agent:	I **understand** your **frustration**, but like other airlines, if the delay is outside our **control** and we have an alternative for you within 5 hours, that's the standard **solution**. But I **know** it's not **fair** to you, so to **try** to do **better**, I can give you a flight bank voucher for the same **amount,** which you can use any time in the next year.
Control	You:	Huh. That doesn't seem ideal to me... But. Ok, fine just give me the credit.
Control	Agent:	So, give me just a moment on that **change**. **Okay**, it's looking **good**.
Control	You:	And do you know why this flight was cancelled?
Control	Agent:	**Apparently,** with the **horrible** weather up there, they **unfortunately** had to cancel all the flights into Chicago this morning.
Control	You:	Perfect storm.
Control	Agent:	**Awful**. **Okay**, you paid $146, and I put that amount in your flight bank, and it can be **used** whenever you **choose** for the next year. I **hope** that sounds **alright**?
Control	You:	Okay.
Control	Agent:	So I **trust** that's all, but was there **anything** else I can **help** with today?
Control	You:	No, that's it.
Control	Agent:	**Alright**, very **well**. and just so you're **aware**, it can take 3-5 days for the credit **activation** to be **completed** in the system.
Control	You:	Okay.
Control	Agent:	**Great**. I **know** it's too **bad** about the weather, but **glad** you'll **make** it there **okay**.
Control	You:	Thank you.
Control	Agent:	You're **welcome**, I'm **glad** we **found** a **resolution**. **Thank** you.

Condition	Speaker	Conversational Language
Cog at Start	Agent:	This is Jet Airline customer **support solutions**. My name is Charlie. I'm **committed** to **solving** your **needs**.
Cog at Start	You:	Hi, I was on a flight that was supposed to get out from here to Chicago this morning at 10:50. And I'm not thrilled it was cancelled.
Cog at Start	Agent:	That's **not** a **good outcome**. If you have a confirmation number, I can **best examine** the **problem** to **inform** you on the **cause.**
Cog at Start	You:	Um, sure. But I don't think you're hearing me…
Cog at Start	Agent:	Yes, I **comprehend**. Your flight time **changed**. I can **help determine** the **correct information** and **perspective** on that for **sure**.
Cog at Start	You:	The confirmation number is J2Y5FZ, but I desperately need to get on a new flight now. I don't need an explanation or information.
Cog at Start	Agent:	Uh, yes I **appreciate** that **insight**. **Please** let me **explain** this **problem**. I'll **need** the name **attributed** to the ticket.
Cog at Start	You:	It's [your name]
Cog at Start	Agent:	**Alright**. One moment **please**. I'm just **trying** to **hopefully find** something **good** for you. OK, I've **found** it. Our system booked you on the next available flight. That's 3:15pm.
Cog at Start	You:	Hilarious. Um, that's not gonna work for me, so I'm actually on the way to the airport now to get a different flight, but I'm hoping you can credit me the amount I spent on this flight.
Cog at Start	Agent:	Let me see if I can take **care** of you... **Okay**, **because** your flight was **regrettably** delayed due to **bad** weather, I can **make** this into a flight credit for you.
Cog at Start	You:	No, I'd rather have a full credit. A refund.
Cog at Start	Agent:	I **understand** your **frustration**, but like other airlines, if the delay is outside our **control** and we have an alternative for you within 5 hours, that's the standard **solution**. But I **know** it's not **fair** to you, so to **try** to do **better**, I can give you a flight bank voucher for the same **amount,** which you can use any time in the next year.
Cog at Start	You:	Huh. That doesn't seem ideal to me... But. Ok, fine just give me the credit.
Cog at Start	Agent:	So, give me just a moment on that **change**. **Okay**, it's looking **good**.
Cog at Start	You:	And do you know why this flight was cancelled?
Cog at Start	Agent:	**Apparently,** with the **horrible** weather up there, they **unfortunately** had to cancel all the flights into Chicago this morning.
Cog at Start	You:	Perfect storm.
Cog at Start	Agent:	**Awful**. **Okay**, you paid $146, and I put that amount in your flight bank, and it can be **used** whenever you **choose** for the next year. I **hope** that sounds **alright**?

Cog at Start	You:	Okay.
Cog at Start	Agent:	So I **trust** that's all, but was there **anything** else I can **help** with today?
Cog at Start	You:	No, that's it.
Cog at Start	Agent:	**Alright**, very **well**. And just so you're **aware**, it can take 3-5 days for the credit **activation** to be **completed** in the system.
Cog at Start	You:	Okay.
Cog at Start	Agent:	**Great**. I **know** it's too **bad** about the weather, but **glad** you'll **make** it there **okay**.
Cog at Start	You:	Thank you.
Cog at Start	Agent:	You're **welcome**, I'm **glad** we **found** a **resolution**. **Thank** you.

Condition	**Speaker**	**Conversational Language**
Cog at End	Agent:	This is Jet Airline customer **support**. My name is Charlie. **How** can I assist you today?
Cog at End	You:	Hi, I was on a flight that was supposed to get out here to Chicago this morning at 10:50. And I'm not thrilled it was cancelled.
Cog at End	Agent:	That's **not good**. If you have a confirmation number, I'll do my **best** to **figure** out the issue.
Cog at End	You:	Um, sure. But I don't think you're hearing me…
Cog at End	Agent:	Yes, I hear you. Your flight time **changed**. I can **help** you with that for **sure**.
Cog at End	You:	The confirmation number is J2Y5FZ, but I desperately need to get on a new flight now. I don't need an explanation or information.
Cog at End	Agent:	Uh, yes I **appreciate** that. **Please** let me address this issue. I'll **need** the name on the ticket.
Cog at End	You:	It's [redacted]
Cog at End	Agent:	**Alright**. One moment **please**. I'm just **trying** to **hopefully find** something **good** for you. OK, I've **found** it. Our system booked you on the next available flight. That's 3:15pm.
Cog at End	You:	Hilarious. Um, that's not obviously gonna work for me, so I'm actually on the way to the airport now to get a different flight, but I'm hoping you can credit me the amount I spent on this flight.
Cog at End	Agent:	Let me see if I can take **care** of you... **Okay**, **because** your flight was **regrettably** delayed due to **bad** weather, I can **make** this into a flight credit for you.

Cog at End	You:	No, I'd rather have a full credit. A refund.
Cog at End	Agent:	I **understand** your **frustration**, but like other airlines, if the delay is outside our **control** and we have an alternative for you within 5 hours, that's the standard **solution**. But I **know** it's not **fair** to you, so to **try** to do **better**, I can give you a flight bank voucher for the same **amount,** which you can use any time in the next year.
Cog at End	You:	Huh. That doesn't seem ideal to me... But. Ok, fine just give me the credit.
Cog at End	Agent:	So, give me just a moment on that **change**. **Okay**, it's looking **good**.
Cog at End	You:	And do you know why this flight was cancelled?
Cog at End	Agent:	**Apparently,** with the **horrible** weather up there, they **unfortunately** had to cancel all the flights into Chicago this morning.
Cog at End	You:	Perfect storm.
Cog at End	Agent:	**Awful**. **Okay**, you paid $146, and I put that amount in your flight bank, and it can be **used** whenever you **choose** for the next year. I **hope** that sounds **alright**?
Cog at End	You:	Okay.
Cog at End	Agent:	So I **trust** that's **everything**, but was there **something** else **perhaps** I can **help** **solve** today?
Cog at End	You:	No, that's it.
Cog at End	Agent:	**Alright**, very **well**. And just so you're **aware**, I **find** it **should** take 3-5 days for the credit **activation** to be **completed** and **appear** in the system.
Cog at End	You:	Okay.
Cog at End	Agent:	**Great,** it's **solved** then**.** I **know** it's too **bad** about the **problematic** weather, but **glad** the **conclusion** is you'll still **make** it there **okay**.
Cog at End	You:	Thank you.
Cog at End	Agent:	**Super**. You're **quite** **welcome**, I'm **genuinely** **glad** we **found** a **resolution**. **Thank** you for your **inquiry**.

(Words from the affective process and cognitive process dictionaries are presented in ***red*** *and* ***blue*** *bold font, respectively.)*

Study 6 Results without the Exclusion

Exclusion. Study 6 used the words-per-minute standard exclusion approach described in the pre-registration (https://aspredicted.org/CQP_GPQ). We aimed to have a final sample size of 135 participants per condition after exclusions for a total target N = 810. To account for the exclusion, we asked Prolific for 932 participants (810 + 15%). Nine hundred forty two participants actually completed the study. After the standard exclusion, 806 participants remained, corresponding to an exclusion rate of 14%. Achieved power after exclusion was 99% (α = .05). Results following the pre-registered exclusion are reported in the main manuscript. As shown below, all results replicate without the exclusion.

Results without Exclusion. In addition to a main effect of severity ($F(1, 936) = 16.87$, $p < .001$) and language ($F(1, 936) = 17.62$, $p < .001$), results revealed a severity x language interaction ($F(2, 936) = 10.78$, $p < .001$). Following the pre-registration, we decomposed this interaction into the two focal subsets of the data for which we had predictions. More cognitive language at start versus control, and more cognitive language at end versus control.

Cognitive Language at Start. Simple effects of language ($F(1, 628) = 11.18$, $p < .001$) and severity ($F(1, 628) = 5.57$, $p = .019$) were qualified by the predicted severity x language interaction ($F(1, 628) = 17.53$, $p < .001$). Consistent with the moderation in the field data, while using more cognitive language at the start decreased customer satisfaction (relative to the control) when the issue was higher severity (M = 4.77, SD = 1.59 vs. M = 5.63, SD = 1.36, $F(1, 628) = 26.87$, $p < .001$, $\eta^2_p = .08$), it had no effect when the issue was lower severity (M = 5.52, SD = 1.42 vs. M = 5.42, SD = 1.32; $F(1, 628) = .38$, $p = .539$, $\eta^2_p = .00$).

Cognitive Language at End. An effect of language ($F(1, 619) = 6.49$, $p = .011$) was qualified by the predicted severity x language interaction ($F(1, 619) = 13.95$, $p < .001$). Consistent with the moderation in the field data, while using more cognitive language at the end increased customer satisfaction (relative to the control) when the issue was lower severity (M = 6.06, SD = .92 vs. M = 5.42, SD = 1.32, $F(1, 619) = 24.98$, $p < .001$, $\eta^2_p = .07$), it had no effect when the issue was higher severity (M = 5.51, SD = 1.46 vs. M = 5.63, SD = 1.36; $F(1, 619) = .58$, $p = .447$, $\eta^2_p = .00$).

WEB APPENDIX F: PERCEPTUAL MECHANISM EXPLORATION

Given our contribution pertains to demonstrating moderation of previously theorized language effects (i.e., related to warmth and competence) by *when* over conversational time they occur, re-assessing the psychological mechanism(s) of such effects was not central to the present research. That said, for two studies (Studies 4A and 5) we pre-registered and collected warmth and competence perceptions as exploratory measures. We report the results here to help inform future work that might delve more deeply into the mechanisms behind dynamic language effects.

A central question is how potential psychological mechanisms should be measured given the effects are dynamic with time. For example, temporal language effects may simply mean shifting the same amount of a feature (e.g., warmth) to a different moment, meaning that overall perceptions of warmth or competence might not always change. Future studies could consider moment-to-moment measures (cf. Ramanathan and McGill 2007), to better assess such temporal shifts.

In this appendix, we report the results of traditional static measurement of social perceptions of employee warmth and competence captured as exploratory mediators in Studies 4A and 5 as a first step towards exploring how temporal variation in a language feature might relate to static perceptual outcomes.

Study 4A Exploratory Perceptual Mechanisms

After collecting the dependent measure described in Study 4A we measured participant perceptions of the employee's warmth and competence ("How warm [competent] was the agent?"; 1 = not at all, 7 = very much). Given these potential perceptual mediators are significantly correlated (r = .28, $p < .001$), we explore them both independently and simultaneously (Pieters 2017).

Simple mediation analysis (PROCESS model 4; Hayes 2018) considering perceived warmth as the driver of the relationship between customer satisfaction and our dynamic language recommendation supports warmth as a mediator (indirect effect = .080, 95% CI [.024, .157]). Using more affective words at the start and end made the agent seem warmer (b = .271, $t = 2.93$, $p = .004$), which increased customer satisfaction (b = .295, $t = 5.30$, $p < .001$). Considering competence perceptions seperately found that it also helped drive the effect, albeit only marginally (indirect effect = .068, 90% CI [.008, .134]). The dynamic language recommendation was marginally linked to competence perceptions (b = .146, t = 1.87, $p = .064$), which was itself linked to customer satisfaction (b = .464, t = 7.63, $p < .001$).

Results were similar for simultaneous parallel mediation analysis (PROCESS model 4; Hayes 2018) by both warmth and competence perceptions. The effect was driven by perceived warmth (indirect effect = .057, 95% CI [.013, .126]) and competence, albeit only marginally for the latter (indirect effect = .059, 90% CI [.007, .119]). Using more affective words at the start and end made the agent seem warmer (b = .281, $t = 3.04$, $p = .003$), which increased customer satisfaction (b = .203, $t = 4.01$, $p < .001$). While competence had a positive relationship with customer satisfaction (b = .403, $t = 6.48$, $p < .001$), our language manipulation only marginally shifted competence perceptions (b = .146, $t = 1.87$, $p = .064$).

For thoroughness, we also report these mediation results for the secondary helpfulness measure used by the Study 1 retailer.

Results. Simple mediation analysis (PROCESS model 4; Hayes 2018) considering perceived warmth as the driver of the relationship between helpfulness and our dynamic language recommendation supports warmth as a mediator, albeit marginally (indirect effect = .035, 90% CI [.004, .073]). Using more affective words at the start and end made the agent seem significantly warmer (b = .254, t = 2.74, p = .007), which increased helpfulness (b = .136, t = 2.01, p = .046). Considering competence perceptions seperately found that it also helped drive the effect, albeit only marginally (indirect effect = .068, 90% CI [.008, .134]). The dynamic language recommendation was marginally linked to competence perceptions (b = .148, t = 1.88, p = .062), which was itself linked to customer satisfaction (b = .486, t = 6.87, p < .001).

Simultaneous parallel mediation (PROCESS model 4; Hayes 2018) by both warmth and competence found that competence marginally mediated the relationship (indirect effect = .071, 90% CI [.001, .136], while warmth was not significant (indirect effect = .008, 90% CI [-.016, .038]). Our dynamic language condition made the agent seem somewhat more competent (b = .148, t = 1.88, p = .062), which increased perceived helpfulness (b = .477, t = 6.48, p < .001). While the dynamic language condition increased perceptions of warmth (b = .265, t = 2.84, p = .005), warmth was not linked to helpfulness (b = .029, t = .47, p = .641).

Discussion. Our exploratory mediation analysis generally suggests the dynamic language recommendation enhanced customer satisfaction because it made the agent seem warmer and/or more competent. Which perception was stronger depended on the outcome measure (customer satisfacation or helpfulness) and the model used (simple vs. parallel mediation). Our preliminary interpretation of these results is that while both warmth and competence should drive customer satisfaction, competent language's effect may have been weaker because it appeared in the conversation's middle in both conditions, which is when our dynamic model suggests this language is likely to shape customer satisfaction.

Why then might competence have been a stronger mediator than warmth when it came to the secondary helpfulness outcome? We speculate that this could have occured because competence perceptions are more clearly linked to assessing whether someone actually helped (i.e., agentically helped solve an issue). To attempt to shed further light on the mechanism(s), Study 5 offers a replication test of the perceptual mechanism(s) through which our dynamic recommendation shapes customer satisfaction.

Study 5 Exploratory Perceptual Mechanisms

Study 5 offers an additional exploratory test of the perceptual mechanism(s) through which our dynamic recommendation shapes customer satisfaction using the same measures as Study 4A. As in Study 4A, we measured participant perceptions of the employee's warmth and competence ("How warm [competent] was the agent?"; 1 = not at all, 7 = very much) after collecting the dependent measure. Because these potential perceptual mediators are significantly correlated (r = .70, $p < .001$), we explore them both independently and simultaneously (Pieters 2017).

Results. Simple mediation analysis (PROCESS model 4; Hayes 2018) considering perceived warmth as the driver of the relationship between customer satisfaction and our dynamic language recommendation supports warmth as a mediator (indirect effect = .251, 95% CI [.102, .417]). Using more affective words at the start and end made the agent seem significantly warmer (b = .333, $t = 3.25$, $p = .001$), which increased satisfaction (b = .756, $t =$

13.00, $p < .001$). Considering competence perceptions seperately found that it also drives the effect (indirect effect = .199, 95% CI [.029, .362]). Our dynamic language recommendation was positively linked to agent competence perceptions (b = .219, t = 2.30, $p = .023$), which was itself linked to customer satisfaction (b = .909, t = 17.16, $p < .001$).

Simultaneous parallel mediation analysis (PROCESS model 4; Hayes 2018) finds that the effect of our dynamic language recommendation on customer satisfaction was driven in parallel by both perceived warmth (indirect effect = .111, 95% CI [.030, .220]) and competence (indirect effect = .145, 95% CI [.025, .268]). Using more affective words at the start and end rather than in the middle made the agent seem warmer (b = .333, $t = 3.25$, $p = .001$), which increased customer satisfaction (b = .334, $t = 5.42$, $p < .001$). Similarly, using more cognitive words in the middle rather than at the start and end made the agent seem more competent (b = .219, $t = 2.30$, $p = .023$), which increased customer satisfaction (b = .664, $t = 9.99$, $p < .001$).

For thoroughness, we also report these mediation results for the secondary helpfulness measure used by the Study 1 retailer.

Simple mediation analysis (PROCESS model 4; Hayes 2018) considering perceived warmth as the driver of the relationship between helpfulness and our dynamic language recommendation supports warmth as a mediator (indirect effect = .177, 95% CI [.070, .290]). Using more affective words at the start and end made the agent seem significantly warmer (b = .333, $t = 3.25$, $p = .001$), which increased perceived helpfulness (b = .532, $t = 9.74$, $p < .001$). Considering competence perceptions seperately found that it also drives the effect (indirect effect = .172, 95% CI [.026, .320]). Our dynamic language recommendation was positively linked to agent competence perceptions (b = .219, t = 2.30, $p = .023$), which was itself linked to helpfulness (b = .787, t = 20.11, $p < .001$).

Simultaneous parallel mediation analysis (PROCESS model 4; Hayes 2018) finds that the effect of our dynamic language recommendation on helpfulness was driven by perceived competence (indirect effect = .162, 95% CI [.049, .280]) but not warmth under this specification (indirect effect = .021, 90% CI [-.015, .060]). Our dynamic language recommendation made the agent seem more competent (b = .333, t = 3.25, p = .001), which increased helpfulness (b = .741, t = 13.88, p < .001).

As in the Study 4A exploratory mediation analysis, the present studies exploration found that both warmth and competence mediated the customer satisfaction outcome. But once again, competence was a stronger mediator than warmth when it came to the secondary helpfulness outcome used by the Study 1 firm. As in Study 4A, we speculate that this might have occurred again here because competence perceptions are more clearly linked to assessing whether someone actually helped (i.e., agentically helped solve an issue).

Study 4A Exploratory Perceptual Mechanisms without the Exclusion

After collecting the dependent measure described in Study 4A we measured participant perceptions of the employee's warmth and competence ("How warm [competent] was the agent?"; 1 = not at all, 7 = very much).

Simple mediation analysis (PROCESS model 4; Hayes 2018) considering perceived warmth as the driver of the relationship between customer satisfaction and our dynamic language recommendation supports warmth as a mediator (indirect effect = .077, 95% CI [.022, .150]). Using more affective words at the start and end made the agent seem warmer (b = .262, t = 3.13, p = .002), which increased customer satisfaction (b = .293, t = 4.79, p < .001). Considering competence perceptions seperately found that it also helped drive the effect (indirect effect = .074, 95% CI [.014, .145]). The dynamic language recommendation was positively linked to competence perceptions (b = .174, t = 2.47, p = .014), which was itself linked to customer satisfaction (b = .426, t = 6.06, p < .001).

Results were similar under simultaneous parallel mediation analysis (PROCESS model 4; Hayes 2018) by both warmth and competence perceptions. The effect was driven by perceived warmth (indirect effect = .057, 95% CI [.012, .121]) and competence (indirect effect = .063, 95% CI [.011, .126]). Using more affective words at the start and end made the agent seem warmer (b = .271, t = 3.24, p = .001), which increased customer satisfaction (b = .207, t = 5.12, p < .001). Competence perceptions also shifted due to our dynamic recommendation (b = .174, t = 2.47, p = .014), and competence perceptions were linked to customer satisfaction (b = .361, t = 5.12, p < .001).

For thoroughness, we also report these mediation results for the secondary helpfulness measure used by the Study 1 retailer.

Simple mediation analysis (PROCESS model 4; Hayes 2018) considering perceived warmth as the driver of the relationship between helpfulness and our dynamic language recommendation supports warmth as a mediator, albeit marginally (indirect effect = .035, 90% CI [.004, .073]). Using more affective words at the start and end made the agent seem significantly warmer (b = .254, t = 2.74, p = .007), which increased helpfulness (b = .136, t = 2.01, p = .046). Considering competence perceptions seperately found that it also helped drive the effect, albeit only marginally (indirect effect = .068, 90% CI [.008, .134]). The dynamic language recommendation was marginally linked to competence perceptions (b = .148, t = 1.88, p = .062), which was itself linked to customer satisfaction (b = .486, t = 6.87, p < .001).

Simultaneous parallel mediation (PROCESS model 4; Hayes 2018) by both warmth and competence found that competence mediated the relationship (indirect effect = .077, 95% CI [.017, .145], while warmth was not significant (indirect effect = .007, 90% CI [-.015, .035]). Our dynamic language condition made the agent seem somewhat more competent (b = .177, $t = 2.50$, $p = .014$), which increased perceived helpfulness (b = .433, $t = 6.14$, $p < .001$). While the dynamic language condition increased perceptions of warmth (b = .256, $t = 3.05$, $p = .003$), warmth was not linked to helpfulness (b = .028, $t = .48$, $p = .635$).

Study 5 Exploratory Perceptual Mechanisms without the Exclusion

Study 5 offers an additional exploratory test of the perceptual mechanism(s) through which our dynamic recommendation shapes customer satisfaction using the same measures as Study 4A. As in Study 4A, we measured participant perceptions of the employee's warmth and competence ("How warm [competent] was the agent?"; 1 = not at all, 7 = very much) after collecting the dependent measure.

Simple mediation analysis (PROCESS model 4; Hayes 2018) considering perceived warmth as the driver of the relationship between customer satisfaction and our dynamic language recommendation supports warmth as a mediator (indirect effect = .268, 95% CI [.138, .401]). Using more affective words at the start and end made the agent seem significantly warmer (b = .354, $t = 3.52$, $p < .001$), which increased satisfaction (b = .755, $t = 13.19$, $p < .001$). Considering competence perceptions seperately found that it also drives the effect (indirect effect = .213, 95% CI [.077, .353]). Our dynamic language recommendation was positively linked to agent competence perceptions (b = .234, t = 2.51, p = .013), which was itself linked to customer satisfaction (b = .909, t = 17.42, p < .001).

Simultaneous parallel mediation analysis (PROCESS model 4; Hayes 2018) finds that the effect of our dynamic language recommendation on customer satisfaction was driven in parallel by both perceived warmth (indirect effect = .119, 95% CI [.048, .208]) and competence (indirect effect = .155, 95% CI [.056, .258]). Using more affective words at the start and end rather than in the middle made the agent seem warmer (b = .354, t = 3.52, p < .001), which increased customer satisfaction (b = .337, t = 5.58, p < .001). Similarly, using more cognitive words in the middle rather than at the start and end made the agent seem more competent (b = .234, t = 2.51, p = .013), which increased customer satisfaction (b = .663, t = 10.19, p < .001).

For thoroughness, we also report these mediation results for the secondary helpfulness measure used by the Study 1 retailer.

Simple mediation analysis (PROCESS model 4; Hayes 2018) considering perceived warmth as the driver of the relationship between helpfulness and our dynamic language recommendation supports warmth as a mediator (indirect effect = .191, 95% CI [.081, .313]). Using more affective words at the start and end made the agent seem significantly warmer (b = .354, t = 3.52, p < .001), which increased perceived helpfulness (b = .540, t = 9.97, p < .001). Considering competence perceptions seperately found that it also drives the effect (indirect effect = .182, 95% CI [.040, .329]). Our dynamic language recommendation was positively linked to agent competence perceptions (b = .234, t = 2.51, p = .013), which was itself linked to helpfulness (b = .776, t = 18.84, p < .001).

Simultaneous parallel mediation analysis (PROCESS model 4; Hayes 2018) finds that the effect of our dynamic language recommendation on helpfulness was driven by perceived competence (indirect effect = .166, 95% CI [.035, .293]) but not warmth under this specification (indirect effect = .033, 90% CI [-.004, .080]). Our dynamic language recommendation made the agent seem more competent (b = .234, $t = 2.51$, $p = .013$), which increased helpfulness (b = .708, $t = 12.70$, $p < .001$).

As in the Study 4A exploratory mediation analysis, the present studies exploration found that both warmth and competence mediated the customer satisfaction outcome. But once again, competence was a stronger mediator than warmth when it came to the secondary helpfulness outcome measure used by the Study 1 firm. As in Study 4A, we speculate that this might have occurred again here because competence perceptions are more clearly linked to assessing whether someone actually helped (i.e., agentically helped solve an issue).

WEB APPENDIX REFERENCES

Chen, Qingxia, Hui Nian, Yuwei Zhu, H. Keipp Talbot, Marie R. Griffin, and Frank E. Harrell Jr (2016), "Too Many Covariates and Too Few Cases? A Comparative Study," *Statistics in Medicine*, 35(25), 4546-4558.

Heinze, Georg, Christine Wallisch, and Daniela Dunkler (2018), "Variable Selection – A Review and Recommendations for the Practicing Statistician," *Biometrical Journal*, 60(3), 431-449.

Meier, Lukas, Sara Van De Geer, and Peter Buhlmann (2008), "The Group Lasso for Logistic Regression," *Journal of the Royal Statistical Society: Series B*, 70, 53-71.

Tibshirani, Robert (1996), "Regression Shrinkage and Selection via the Lasso," *Journal of the Royal Statistical Society, Series B*, 58(1), 267-288.

Just, Marcel Adam and Patricia Ann Carpenter (1987). *The Psychology of Reading and Language Comprehension*. Boston: Allyn & Bacon.

Pieters, Rik (2017), "Meaningful Mediation Analysis: Plausible Causal Inference and Informative Communication," *Journal of Consumer Research*, 44(3), 692-716.

Ramanathan, Suresh and Ann L. McGill (2007), "Consuming with Others: Social Influences on Moment-to-Moment and Retrospective Evaluations of an Experience," *Journal of Consumer Research*, 34(4), 506-524.

Yang, Yi and Hui Zou (2015), "A Fast Unified Algorithm for Solving Group-Lasso Penalized Learning Problems," *Statistics and Computing*, 25, 1129-1141.

Yuan, Ming and Yi Lin (2006), "Model Selection and Estimation in Regression with Grouped Variables," *Journal of the Royal Statistical Society, Series B*, 68(1), 49-67.

Zou, Hui (2006), "The Adaptive Lasso and Its Oracle Properties," *Journal of the American Statistical Association*, 101(476), 1418-1429