

Insights from Customers' Chats with Bots and Human Agents
Forthcoming in *Marketing Intelligence and Planning*

Vivek Astvansh*

Associate Professor of Quantitative Marketing and Analytics,
Desautels Faculty of Management, McGill University, Canada
vivek.astvansh@mcgill.ca

Academic Director, Bensadoun School of Retail Management, McGill University, Canada

Adjunct Associate Professor of Data Science, Department of Informatics,
Luddy School of Informatics, Computing, and Engineering,
Indiana University Bloomington, USA

Affiliate, Environmental Resilience Institute,
Indiana University

**Acknowledgments: Vivek wrote this manuscript while he was the Research Director of the Indiana University Bloomington Kelley School of Business' Center for Education and Research in Retail. He thanks the Center's Director, John Talbott, for (1) introducing him to the retailer who provided the data and (2) paying Gautam Chauhan (a master's student at IU's Luddy School), who assisted Vivek in the research.*

Insights from Customers' Chats with Bots and Human Agents

ABSTRACT

Customers often use the “chat” feature on a company’s website to interact with a human agent or a bot and receive customer service. However, companies lack knowledge of (1) why customers use the chat feature, (2) who—the agent or the bot—is more similar (in content) to the customer, and (3) whether and how this similarity impacts the customer’s engagement during the chat. This article provides this knowledge. I partnered with a U.S. public retailer to receive transcripts of nearly 500,000 chats between a customer and an agent, a bot, or both. I conducted three analyses. Using KeyBERT and a guided LDA, the first analysis reveals that customers prefer chatting with an agent (vs. the bot) when they seek detailed or sensitive information. Analysis 2 uses three similarity measures to demonstrate that relative to the bot, the agent is more similar (in content) to the customer. Analysis 3 uses guided LDA and gradient boosting (XGBoost) to show that matching the customer on the dominant topic boosts customer engagement during the chat. The findings help retail managers design better chat features and chatbots, thus improving customer engagement.

Keywords: customer service, chatbot, text analysis, guided LDA, KeyBERT, embedding

INTRODUCTION

Companies have long provided customers a “human agent” chat feature on their website, app, or social media pages. Recently, companies have introduced a “*bot agent*” in this chat feature (Davenport et al. 2020). A *chatbot* is a computer program that uses natural language to converse with human users (Holtgraves et al. 2007; Shawar and Atwell 2005). The bot aims to provide the first line of customer service, transferring customers to a human agent (hereafter, agent, for brevity) if and when the bot cannot provide the information the customer is asking for or the customer explicitly asks to be connected to an agent. Industry reports forecast that by 2025, 95% of all customer interactions will be augmented or replaced by chatbots (Servion 2020). By some estimates, the chatbot market is expected to reach \$5 billion by 2023 (Fokina 2024). Although prior research has examined customer-bot interactions (Crolic, Thomaz, Hadi, and Stephen 2021; Kull, Romero, and Monahan 2021; Mariani, Hashemi, and Wirtz 2023; Mozafari, Weigner, and Hammerschmidt 2022; Schanke, Burtch, and Ray 2021), academics and managers lack empirical evidence on why customers may prefer to chat with a bot (vs. a live agent) and whether chatting with a bot boosts customer engagement.

How pervasive is the chat feature—and the chatbot in particular—and how well do companies know why customers use a chatbot? I checked the websites of the top 25¹ retailers in the United States. Of these top 25 retailers, 16 provide a chat feature, and 14 include a bot as first-line support. More importantly, some retailers did not offer any prepopulated list of reasons why a customer should chat with a bot. In contrast, others provided a rudimentary list of reasons, such as order status or canceling an order. This preliminary evidence guides my first research question: *Why do customers use the chat feature?* The answer to this question will help managers

¹ As per <https://nrf.com/resources/top-retailers/top-100-retailers/top-100-retailers-2021-list>.

prioritize their resource allocation to more common reasons instead of proceeding with little focus.

Once companies know why customers use the chat feature and provide a bot and an agent, they need to know who is more effective in chatting with a customer—the agent or the bot. Prior research (Herhausen, Ludwig, Grewal, Wulf, and Schoegel 2019; Ireland and Pennebaker 2010; Ludwig, De Ruyter, Friedman, Brügger, Wetzels, and Pfann 2013; Moore and McFerran 2017) has positioned *linguistic style matching* between two conversation partners as a key determinant of partners' engagement. Therefore, I measure the linguistic similarity between the operator's² and customer's messages. Similarity breeds liking (Wooley and Fishbach 1983), and I reason that the more similar the operator is to the agent, the more effective the conversation is. Once one knows the content-level similarity between the two writers (the customer on one side and the operator on the other), one can determine whether such similarity boosts customer engagement during the chat (Berger, Kim, and Meyer 2021). Thus, the research question is: *Who—the agent or the bot—is, on average, more similar (in content) to the customer? And does this similarity boost customer engagement?*

I partnered with a U.S.-based, NASDAQ-listed retailer, and I received transcripts of nearly 500,000 chats between customers and agents, bots, or both from my partner.³ The corpus includes 40,069 customer-agent chats, 341,879 customer-bot chats, and 96,833 customer-bot-agent chats. For each chat, I separated the messages by the writer. Next, I cleaned and lemmatized the messages by each writer and used the resulting “processed” messages to conduct three analyses that answered my three questions. I conduct three analyses on messages sent by

² “Operator” could be the agent or the bot or both (for chats that the bot transfers to an agent).

³ The partner retailer uses a basic chatbot purchased from Salesforce.

customers, bot (agent), and (human) agents. Analyses 1 and 2 operate at the corpus level, whereas Analysis 3 is at the chat level.

Analysis 1 reveals two insights from the topics of messages written by these three writers. First, customers often use the chat feature to receive order-related information, followed by information about coupons and products. Second, customers prefer the bot to transfer them to an agent for two topics—shipping and payment—indicating an aversion to chatting with a bot when the customer wants detailed or sensitive information, as in the case of shipping and payment, respectively. Analysis 2 answers who—the agent or the bot—has a linguistic style like the customer’s. It reveals that the agent is more like the customer than the bot is. The finding suggests that human agents are superior to state-of-the-art chatbots in demonstrating similarity to the customer in the content. Analysis 3 builds on Analyses 1 and 2 by asking whether a content-level match between the customer and the operator (which could be the bot or the [human] agent) boosts the customer’s engagement during the chat (Packard and Berger 2021). I build three XGBoost models to predict customer engagement. The results suggest that including the similarity in an operator’s dominant topic and the customer’s topic boosts the model’s prediction of customer engagement.

My research contributes to two streams of literature. First, it adds to the emerging knowledge of customer-bot interactions. This research has documented customer responses to chatbot humanization (e.g., Crolic et al. 2022; Schanke, Burtch, and Ray 2021) and language use (e.g., Jiang et al. 2022; Yun and Park 2021). I contribute by listing why customers prefer chatting with a bot over a human agent. In addition, I show that service agents’ conversational style is more similar (than the bot’s) to customers’ style. This similarity boosts customer engagement, highlighting human agents’ role in serving customers (Gnewuch, Morana, Hinz, Kellner, and

Maedche 2024). The finding also extends prior knowledge on linguistic style matching between conversation partners (Herhausen, Ludwig, Grewal, Wulf, and Schoegel 2019; Ireland and Pennebaker 2010; Ludwig, De Ruyter, Friedman, Brügger, Wetzels, and Pfann 2013; Moore and McFerran 2017).

Second, I contribute to the emerging literature on drawing marketing insights from customer- and marketer-generated text. This literature has focused on such text's *persuasiveness* (Berger and Packard 2022; Humphreys and Wang 2017; Kronrod 2022; Packard and Berger 2024; Pogacar, Shrum, and Lowrey 2017, 2018). I use the full back-and-forth conversation (Shawar and Atwell 2005) between a customer and a company representative (both agent and bot) instead of a one-off unidirectional message from a customer to a company (as is the case in user-generated content on social media platforms). The resulting analysis allows me to extend research by using customer-agent and customer-bot conversations (Gnewuch, Morana, Hinz, Kellner, and Maedche 2024).

Retail managers can use my findings to focus on the top three reasons and improve the chat feature's effectiveness in delivering value to the customer on these features. Such focused improvement can yield more favorable outcomes than improving customer service on multiple dimensions. Further, the finding that the bot uses more words and a broader vocabulary than the human agent should surprise technology companies that develop chatbots, such as Salesforce, Microsoft, Google, and Amazon. These companies should program their chatbots to be more efficient (i.e., use fewer words) and effective (i.e., use precise words), thus offering customers precise and consistent communication. Relatedly, my finding that bots are less like human agents in mimicking customers should lead these companies to boost their chatbots' learning from the

customer's prior language use, thus boosting customer engagement and firm outcomes arising from such engagement.

LITERATURE ON CUSTOMER-CHATBOT INTERACTIONS

I organize the existing literature⁴ into two groups: (1) customer response to a firm's *disclosure* that they are chatting with a bot and not a human agent, and (1) customer response to chatting with a bot (vs. a human agent) (see Mariani, Hashemi, and Wirtz's [2023] review article).

Using a field experiment, Luo, Tong, Fang, and Qu (2019) showed that an undisclosed text-based chatbot elicited as many customer purchases as experienced agents. Importantly, the chatbot was four times more effective than an inexperienced agent. However, the firm's disclosure that they were chatting with a machine reduced the customer's purchase likelihood by nearly 80%. Similarly, Gnewuch, Morana, Hinz, Kellner, and Maedche (2024) used field and lab experiments to show that disclosing a human agent's involvement in a customer's conversation with a bot leads the customer to adopt a human-like style of conversation, triggered by their need to manage impressions. Such adoption increases employee workload because many customer requests get routed to the human agent.

The second stream of research has examined customer response to service offered by a chatbot versus a (human) agent. Broadly, academics have reported that customers believe that a firm uses a chatbot to cut costs at the expense of providing customers with superior service. This belief makes customers evaluate their bot interactions more negatively than agent interactions.

⁴ I acknowledge the literature on determinants of a customer's chatbot adoption (Mariani, Hashemi, and Wirtz 2023; Pentina, Xie, Hancock, and Bailey 2023). This literature has drawn on the broader literature on the technology acceptance model (Davis 1989), which posits that (1) the easier a user perceives to use a technological product and (2) the greater their perceived benefits from the use, the more likely they are to adopt the product. I also acknowledge the broader literature on the use of technology in facilitating customer-firm interactions (see Grewal, Hulland, Kopalle, and Karahanna 2020; Hoffman, Moreau, Stremersch, and Wedel 2022; Huang and Rust 2018, 2024; Liu-Thompkins, Okazaki, and Li 2022). I summarize the literature that is more relevant to my article.

Consistent with the belief, the effect is (1) nullified if the firm discounts the price, thus sharing the economic surplus with them, and (2) reversed if the bot offers superior service than their human counterparts (Castelo, Boegershausen, Hildebrand, and Henkel 2023). Relatedly, academics have examined customer response to a humanized (vs. non-humanized) chatbot. Crolic, Thomaz, Hadi, and Stephen (2021) considered the common customer service context when a customer is angry. They reported that when the customer is angry, a humanized chatbot (i.e., one with a human name and an avatar) yields lower customer satisfaction, firm evaluation, and purchase intentions than its non-humanized counterpart. Similarly, Schanke, Burtch, and Ray (2021) found that humanizing a chatbot (by including one or more of three features: humor use, communication delays, and social presence⁵) raises customer conversion rate. However, humanizing a chatbot also increases customers' price sensitivity because they are likelier to negotiate with a humanlike (vs. non-humanized) chatbot.

The preceding summary highlights three gaps. First, academics and managers know little about the topics for which customers prefer a bot over a human agent and vice versa. The answer will help managers better allocate their resources based on customer needs. My Analysis 1 provides this answer. Second, one would expect a bot to be programmed to mimic the customer's linguistic style. However, psychologists have shown that some people are particularly adept at subliminally mimicking others in their use of language (Niederhoffer and Pennebaker 2002). Academics and managers would thus benefit from knowing who is more adept at linguistic mimicking—the bot or the agent. My Analysis 3 provides this knowledge. Third, does linguistic similarity help the firm's outcomes? My chat-level analysis provides an affirmative answer, thus boosting academics' and managers' knowledge of the returns to linguistic mimicking.

⁵ Social presence is the degree to which a chatbot communicates like a person (Short et al. 1976). Examples include human name, use of informal language, and typing cues.

DATA

Empirical Context

I partnered with a U.S.-headquartered NASDAQ-listed retailer. It operates nearly 400 stores in the Midwest, South, and Southeast regions, and is a leader in the family footwear segment. Founded in 1978, it earned sales revenue of \$1.176 billion in fiscal year 2023. It sells men's, women's, children's, and athletic footwear through its brick-and-mortar and electronic stores.

A customer survey revealed that the retailer will boost customer satisfaction by letting customers chat with an agent. Unlike the call center that requires customers to be in a private space and use voice, a chat feature could allow them to interact with the retailer on the go, without requiring private space. Therefore, the retailer introduced *chat with a live agent* on March 12, 2018, and a *basic* chatbot—purchased from Salesforce—on December 14, 2020. *Basic* means that the customer could do naïve question-and-answer with the chatbot. The customer could ask the chatbot to transfer them to a human agent if the chatbot could not understand the customer's text or if the customer prefers chatting with a human agent.

Raw Data, Exclusion Criteria, and Separation into Discrete Messages

The retailer provided me with transcripts of all 545,987 chats between September 15, 2020, and August 6, 2021. The retailer chose the start date of September 15, 2020, because the date marked the retailer's transition to a new customer service provider. The retailer manages all data in Salesforce Service Cloud. The data file that my counterpart at the retailer exported from Salesforce Service Cloud and provided me includes three key fields: *Chat ID*, *Body*, and *Is Chatbot Session*. *Chat ID* is an 18-character alphanumeric value that uniquely identifies each

chat session. The *Body* is the raw transcript with HTML tags (see Appendix A). *Is Chatbot Session* is an indicator field, which equals 1 if the chat involved the bot and 0, otherwise.

I dropped three types of observations from the full data set of 545,987. First, I excluded 60,862 (11.1%) observations that had an empty *Body*. I consulted my contact person at the retailer to understand the reason for the empty *Body*. The person told me that Salesforce Service Cloud provides a blank value if a customer closes the chat application before it loads in the web browser or the app. This step reduced my data set to 485,125 observations. Second, of the resulting 485,125 chats, 438,712 (90.4%) involved the chatbot (i.e., *Is Chatbot Session* was 1), and the remaining 46,413 (9.6%) did not. Of the 46,413 chats that did not involve a bot—that is, involved only a human agent—203 observations had *Body* = “Test Agent,” indicating that these chats were created for testing. I dropped these 203 observations, ending with 484,922 chats, which comprise 438,712 customer-bot chats and 46,210 customer-agent chats. Third, I wrote a regular expression (“regex”) that separated each transcript into discrete messages such that each message accompanies the writer’s identity and the written text.

```
\s(?:P<Author>[a-zA-Z]+|(?:[A-Za-z]+\s){0,2}[A-Za-z]+\s):(?:P<Message>(?:s).*?(?:\s[0-9]+\s?\s\)|.*?(?:\s[0-9]+m\s[0-9]+\s?\s\)))
```

The regular expression command could not find the specified pattern for 6,141 of the 46,210 customer-agent chats. I read these chats manually and discovered that the *Body* included only HTML content and no message from the customer or the agent. That is, the *Body* variable, for example, had the following value for such 6,141 chats: ‘<p align="center">Chat Started: Friday, August 06, 2021, 15:31:07 (+0000)</p>’. I excluded these 6,141 observations to end up with 478,781 chats, of which 438,712 (91.63%) are customer-bot chats and 40,069(8.37%) are customer-agent chats.

The *customer-agent chats*—by definition—do not involve the bot. In contrast, the *customer-bot chats* mean the customer started the conversation with the bot. However, the bot may have transferred the customer to an agent. Therefore, I separated the 438,712 customer-bot chats into the “correctly identified” customer-bot chats (where the customer chatted with the bot *only*) and customer-bot-agent chats where the bot transferred the chat to an agent. My counterpart at the retailer told me that if a chat includes the message “Hello, how can I assist you today?” this message marks the transition of the chat from the bot to the agent. This default message helped me separate the customer-agent part of a chat from the preceding customer-bot part. Of the 438,712 chats, 341,879 (77.93%) are customer-bot chats, and 96,833 (22.07%) are customer-bot-agent chats (Table 1).

Web Appendix A provides an example raw transcript for each of the three types of chats: customer-agent, customer-bot, and customer-bot-agent.

Next, I re-ran the regex command to separate each *Body* (i.e., chat transcript) into discrete messages written by each writer. The 40,069 customer-agent chats yielded 494,268 messages, of which the customer wrote 40% and 60% by the agent. The 341,879 customer-bot chats included 2,849,469 messages, of which 22.23% were messages written by the customer and 77.77% by the bot. Lastly, the 96,833 customer-bot-agent chats included 5,162,005 messages, such that 28% were written by the customer, 61% by the bot, and 10% by the agent.

Table 1: Distribution of Chats by the Three Types

| Type of Chat | Count | #Messages by all three writers | Distribution of #messages by writer |
|----------------|------------------|--------------------------------|-------------------------------------|
| Customer-agent | 40,069 (8.37%) | 494,268 | 40% by customer, 60% by agent |
| Customer-bot | 341,879 (71.41%) | 2,849,469 | 22.23% by customer, 77.77% by bot |

| | | | |
|--------------------|-----------------|-----------|---|
| Customer-bot-agent | 96,833 (20.22%) | 5,162,005 | 28% by customer, 61% by bot, 10% by agent |
| All chats | 478,781 | 100% | |

Descriptive Statistics

I report three managerially consequential descriptive statistics. First, I calculated the number of words sent by each writer in each chat. On average, a customer wrote 18.81 words, an agent 20.18, and a bot 64.89. The insight is that the bot wrote more than an agent or a customer. Alternatively, the bot made the customer read far more text than the agent. Similarly, the bot typed nearly four messages per session. In contrast, the agent wrote 1.23, and the customer .95. Managers may consider making the bot more efficient because customers may prefer less text and more efficiency.

Second, I calculated the Type-Token-Ratio (TTR), which helped me determine the uniqueness of the words in each writer's message in each chat.

$$TTR = \frac{\# \text{ of unique words in a message}}{\# \text{ of words in a message}}$$

The average uniqueness (TTR) in a customer's messages is 54.04, whereas that in the agent's and bot's messages is 22.47 and 70, respectively. High uniqueness in customer messages is expected because customers are heterogeneous and thus vary in their vocabulary. Low uniqueness in an agent's messages is also expected because agents are trained to use a more condensed vocabulary. What surprises me is that the bot's vocabulary includes a larger set of words. Managers may consider limiting the bot's vocabulary to enable the bot's precise and predictable use of words while conversing with the customer.

Third, because customers use the chat feature to ask questions, I detected question marks in each writer's message in each chat. On average, the bot asked 1.8 questions, the agent asked

.6, and the customer asked .49. This finding is consistent with my expectations. A bot is better programmed to ask questions and frame them as questions. In contrast, an agent may ask questions but often forget to frame them as questions.

Preprocessing

At this point, I had separated the *Body* of each of the 478,781 chats into discrete messages by the writer. I had three fields for each *Chat ID*: *Raw Customer Messages*, *Raw Agent Messages*, and *Raw Bot Messages*. Each of these three fields is a concatenated string of the individual messages that the focal writer wrote, where I used the pipe character to separate the messages.

While *Raw Customer Messages* has a value for each chat, *Raw Agent Messages* is blank for customer-bot chats (which, by definition, do not include any message from the agent), and *Raw Bot Messages* is empty for customer-agent chats. *Raw Agent Messages* and *Raw Bot Messages* have values for customer-bot-agent chats.

Next, I dropped punctuation marks and stop-words for these three fields. Specifically, I excluded the words listed in the stop-word list of the *nlTK* open-source library, days of a week, months of a year, and noninformative words such as button, and align, which I identified by manual reading of the chat transcript. My stop word list includes 182 unigrams. Removing punctuation marks and stop words led me to a clean copy of the three fields: *Clean Customer Messages*, *Clean Agent Messages*, and *Clean Bot Messages*. Next, I lemmatized—that is, converting each word into its base or dictionary form, called a lemma—each of the three cleaned messages. This step produced values of *Lemmatized Customer Messages*, *Lemmatized Agent Messages*, and *Lemmatized Bot Messages*. I next describe the three analyses I performed on the processed data (Table 2). Tables A2A, A2B, and A2C in the Web Appendix provide examples of raw, cleaned, and lemmatized messages for customer, agent, and bot, respectively.

Table 2: Fields in the Preprocessed Data Set

| Field name | Field description |
|-------------------------------------|---|
| <i>Body</i> | Raw text transcript |
| <i>Raw Customer Messages</i> | Unclean/raw messages sent by the customer. A pipe character separates the messages. |
| <i>Raw Agent Messages</i> | Unclean/raw messages sent by the agent. A pipe character separates the messages. If the chat did not involve an agent, this field is empty. |
| <i>Raw Bot Messages</i> | Unclean/raw messages sent by the bot. A pipe character separates the messages. If the chat did not involve an agent, this field is empty. |
| <i>Clean Customer Messages</i> | Clean/raw messages sent by the customer. A pipe character separates the messages. Clean = Unclean – punctuation marks – stop words |
| <i>Clean Agent Messages</i> | Clean/raw messages sent by the agent. A pipe character separates the messages. If the chat did not involve an agent, this field is empty. |
| <i>Clean Bot Messages</i> | Unclean/raw messages sent by the bot. A pipe character separates the messages. If the chat did not involve an agent, this field is empty. |
| <i>Lemmatized Customer Messages</i> | Lemmatized messages sent by the customer. A pipe character separates the messages. Lemmatized message = Clean message, but with each word replaced with its lemma |
| <i>Lemmatized Agent Messages</i> | Lemmatized messages sent by the agent. The messages are separated by a pipe character. If the chat did not involve an agent, this field is empty. |
| <i>Lemmatized Bot Messages</i> | Lemmatized messages sent by the bot. The messages are separated by a pipe character. If the chat did not involve an agent, this field is empty. |

ANALYSIS 1: KEYWORD EXTRACTION AND TOPIC MODELING

Following recent research in the marketing discipline (Dzyabura and Peres 2021; Peng et al. 2020; Toubia et al. 2018), I used the guided LDA in place of the conventional LDA (i.e., the “unguided” one). As the name suggests, the guided LDA is like supervised machine learning. It involves the data analyst “seeding” the LDA with key terms. Such seeding directs the LDA to look for terms in the vicinity of the key terms instead of exploring all terms in the entire corpus (what the traditional LDA does).⁶

I used *KeyBERT* to create my list of seed key terms to input into the guided LDA.

KeyBERT is a keyword extraction technique that uses BERT embeddings. It operates in three

⁶ I also modeled the data using the traditional LDA and am happy to share the results.

sequential steps. First, it uses BERT to extract embeddings from each “document” (each of the three clean messages in our case) to obtain a document-level representation. Second, it extracts from embeddings “ngrams” (that is, either unigram or bigram). Third, it uses cosine similarity to identify ngrams that best describe the data.

I inputted the cleaned messages for the customer, the agent, and the bot to KeyBERT—one at a time—and programmed *KeyBERT* to produce the topmost unigram or the topmost bigram in the input message.

Keyword Extraction and Guided LDA on Customer Messages

I ran *KeyBERT* on *Clean Customer Messages* of each of the 478,781 chats. This procedure produced the field *Customer Messages – Keywords*. I read this field’s values and identified six intuitive and mutually exclusive topics. Table 3A lists the labels that I provided and the seed words that constitute each of the six topics.

Table 3A: Topics and Seed words for *Customer Messages – Clean*

| Topic | Seed Words |
|-------------|---|
| #1 Order | order status wrong change online cancel miss |
| #2 Coupon | apply coupon certificate perk redeem gift card discount reward code |
| #3 Product | item shoe size boots pair sandal sneaker |
| #4 Shipping | ship free deliver package arrive |
| #5 Account | email unsubscribe reset password |
| #6 Payment | billing visa address payment paypal |

Next, I built a guided LDA on *Clean Customer Messages*, asking for six topics and seeding/guiding the LDA to use the specified key terms (contained in *Customer Messages – Keywords*) for each of the six topics (see Table 3B).

Table 3B: Guided LDA Results on *Clean Customer Messages*, Using *Customer Messages – Keywords* to Guide the LDA

| Topic # and Name | Top unigrams |
|------------------|--|
| #1 Order | order, status, wrong, change, online, cancel, miss |
| #2 Coupon | apply, coupon, certificate, perk, redeem, gift, card, discount, reward |

| | |
|-------------|---|
| #3 Product | item, shoe, size, boots, pair, sandal, exchange |
| #4 Shipping | ship, track, deliver, shipment, pending |
| #5 Account | email, unsubscribe, reset, password |
| #6 Payment | payment, check, error, billing, online, discount, charge, twice |

Keyword Extraction and Guided LDA on Agent Messages

Of the 478,781 chats, 136,902 include an agent (Table 1). Therefore, I ran KeyBERT on *Clean Agent Messages* of the 136,902 chats to collect the key terms in *Agent Messages – Keywords*. I read the values of this field and identified seven topics. Table 4A lists the labels I provided and the seed words constituting the seven topics. The first six topics are the same as those from the customer messages. However, the agent messages include a seventh topic: agent-specific words.

Table 4A: Topics and Seed words for *Agent Messages – Clean*

| Topic # and Name | Seed Words |
|------------------|--|
| #1 Order | Order status wrong change online cancel miss |
| #2 Coupon | Apply coupon certificate perk redeem gift card discount reward |
| #3 Product | Item shoe size boot pair sandal exchange |
| #4 Shipping | Ship track deliver shipment pending |
| #5 Account | Email unsubscribe reset password |
| #6 Payment | Bill visa address payment paypal refund process Klarna |

Next, I built a guided LDA on *Clean Agent Messages*, asking for seven topics and seeding/guiding the LDA to use the specified keywords (contained in *Agent Messages – Keywords*) for each of the seven topics (read Table 4B).

Table 4B: Guided LDA Results on *Clean Agent Messages*, Using *Agent Messages – Keywords* to Guide the LDA

| Topic # and Name | Top unigrams |
|------------------|--|
| #1 Order | Order today number cancel email exchange check ready |
| #2 Coupon | Shoe contact carnival size exchange perk available item |
| #3 Product | Order fedex email today password shoe reset return origin |
| #4 Shipping | Fedex track deliver ship shipment order date receive pending |

| | |
|------------|--|
| #5 Account | Shoe connect help perk sorry email final unfortunately fail hour |
| #6 Payment | Hello assist live discount visa billing issue |

Keyword Extraction and Guided LDA on Bot Messages

I ran *KeyBERT* on *Clean Bot Messages* of each of the 438,712 chats that involved a bot (read Table 1). This procedure produced the field *Bot Messages – Keywords*. I read the values of this field and identified five topics (Table 5A). The first four topics are the same as those from the customer and agent messages. However, the bot messages also include a fifth topic, which I label chat transfer. Further, the bot messages do not include topics of payment and agent, which are featured in agent messages.

Table 5A: Topics and Seed words for *Bot Messages – Clean*

| Topic # and Name | Seed Words |
|------------------|---|
| #1 Order | Order status change online check number pending |
| #2 Coupon | Perks redeem gift |
| #3 Item | Item shoe exchange |
| #4 Chat Transfer | Email unsubscribe |
| #5 Email | Transfer successfully fail business hour sorry |

Next, I built a guided LDA on *Clean Bot Messages*, seeding it to use the specified keywords (contained in *Agent Messages – Clean – Keywords*) and instructing it to produce five topics.

Table 5B: Guided LDA Results on *Clean Bot Messages*, Using *Bot Messages – Keywords* to Guide the LDA

| Topic # and Name | Top unigrams |
|------------------|---|
| #1 Order | 'order', 'status', 'change', 'online', 'check', 'number', 'pending' |
| #2 Coupon | 'perks', 'redeem', 'gift' |
| #3 Item | 'item', 'shoe', 'exchange' |
| #4 Chat Transfer | 'transfer', 'successfully', 'fail', 'business', 'hours', 'sorry' |
| #5 Email | 'email', 'unsubscribe' |

Insight

I note that, unlike the agent messages, the bot messages do not include the topics of shipping and payment. The insight is that the customer converses about these two topics with exclusively the agent and not with the bot. This finding suggests that the customer prefers an agent when they seek detailed information (e.g., shipping) or must converse about sensitive information.

ANALYSIS 2: SIMILARITY IN MESSAGES BY THE THREE WRITERS

Prior research has theorized that linguistic style matching between two conversation partners boosts engagement (Herhausen, Ludwig, Grewal, Wulf, and Schoegel 2019; Ireland and Pennebaker 2010; Ludwig, De Ruyter, Friedman, Brügggen, Wetzels, and Pfann 2013; Moore and McFerran 2017). Therefore, Analysis 2 aims to determine whether the agent or the bot is more like the customer in the content of the messages, while Analysis 3 relates the similarity to customer engagement.

I consider the 96,833 customer-bot-agent chats. I used three measures of similarity: (1) Jaccard similarity index, (2) Levenshtein Distance, and (3) cosine similarity on BERT embeddings. I used the lemmatized versions of customer, agent, and bot messages.⁷

Jaccard similarity index: It compares the two messages at the word level. For example, “AI is our friend and it has been friendly” and “AI and humans have always been friendly” have a total of 10 words, five of which are common. Therefore, the Jaccard similarity index for the two sentences will be $5 \div (5 + 3 + 2) = 0.5$ or 50%.

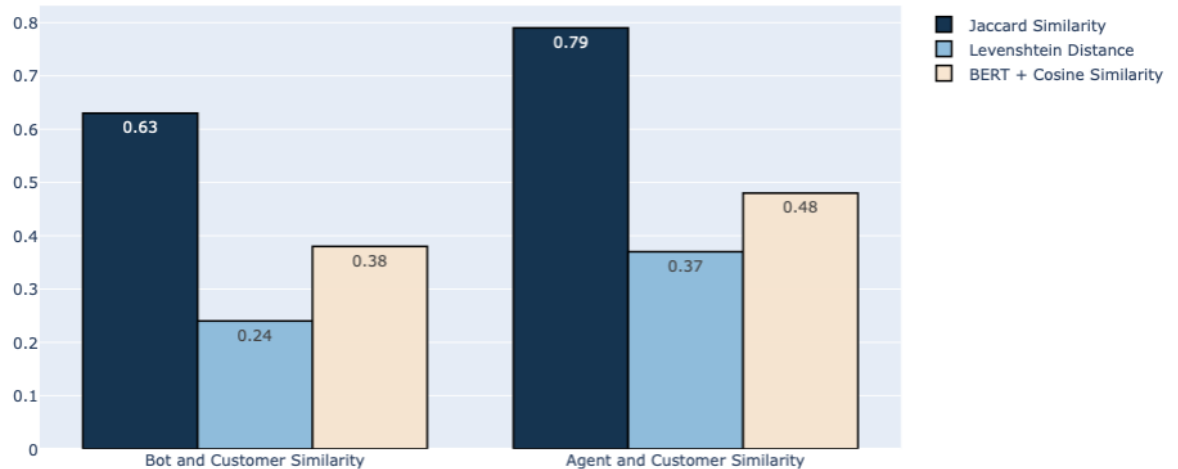
⁷ Calculating similarity based on lemmatized (as opposed to nonlemmatized) messages yields more accurate results. For example, suppose message 1 includes “transferred,” whereas message 2 includes “transferring.” If I compute similarity on nonlemmatized versions of these two messages, they will be dissimilar on the “transfer” words. However, if I reduce the words to their base form “transfer,” the two messages would be similar (<https://ieeexplore.ieee.org/abstract/document/9194665>).

Levenshtein distance: I used the open-source Python library Fuzzywuzzy to compare the strings at the character level (and not word level, unlike the Jaccard Similarity Index). Specifically, the algorithm counts the number of characters to be inserted, deleted, or substituted to make one string the same as the other. It outputs a score between 0 and 100, with 100 meaning that the two strings are the same.

BERT embeddings + cosine similarity: I proceeded in two steps to create data for my third measure of distance. First, I used a pre-trained BERT model designed for semantic search. It has been trained on 215 million question-and-answer pairs from diverse sources and can understand the context. I use this model to compute embeddings—that is, convert text into numbers. Second, for each of the 96,833 customer-bot-agent chats, I computed the cosine similarity between messages written by the customer and the bot and those written by the customer and the agent.

Figure 9 shows the three similarity measures. The key finding is that the agent is more similar (than the bot) to the customer, regardless of which distance measure I consider.

Figure 1: Mean Similarity Between the Corpus of Customer Messages and Bot Messages, and the Corpus of Customer Messages and Agent Messages



ANALYSIS 3: HOW SIMILAR IN CONTENT ARE AN AGENT AND THE BOT TO THE CUSTOMER?

Next, I aim to determine, for each chat, whether the two parties in the conversation—the customer and the agent/bot—are talking about the same topic. I hypothesize that if the two parties match on the dominant topic, the customer will be more engaged. The intuition is that if two parties talk, the conversation may include more than one topic. However, if the two parties focus on the same topic—while diverging on other topics—the conversation will be engaging.

Further, for chats where the two parties match on the dominant topic, the difference in the weights each party gives to the dominant topic will further determine the customer's level of engagement. I hypothesize that the lower this difference, the more engaged the customer will be. The underlying logic is that while matching on the dominant topic is necessary, it may not be sufficient for engagement. Sufficiency comes from paying equal emphasis on the dominant topic.

I test this hypothesis by measuring engagement by the average number of seconds the customer took between consecutive messages.

Method

I am interested in measuring whether the two conversation partners in a chat match on the dominant topic and, if yes, the difference between the weights (i.e., the proportion of words) they give to the dominant topic. Therefore, I separated my data set of 478,781 chats into two data sets. Next, for an apples-to-apples comparison between the two data sets, I selected three topics that are common and pervasive among all three parties.

Table 6: Topics and Seed Words for Topic Similarity

| Topic # and Name | Seed Words |
|-------------------------|---|
| #1 Order and Tracking | 'order', 'status', 'wrong', 'change', 'online', 'cancel', 'missing', 'item', 'shoes', 'size', 'shoe', 'boots', 'pair', 'sandals', 'exchange', 'shipping', 'tracking', 'delivered', 'shipment', 'shipped', 'pending' |
| #2 Account | 'email', 'unsubscribe', 'reset', 'password', 'account' |
| #3 Payment | 'billing', 'visa', 'address', 'payment', 'paypal', 'refund', 'processed', 'klarna', 'apply', 'coupon', 'certificate', 'perk', 'redeem', 'gift', 'cards', 'discount', 'rewards', 'balance' |

Of the 129,655 customer-agent chats, 75,034 (57.8%) have the same dominant topic for the customer and the agent. Of the 259,259 customer-bot chats, 111,397 (42.9%) have the same dominant topic for the customer and the bot.

Next, for the 75,034 chats where the customer and the agent matched on the dominant topic, I computed the difference in the weights each party gave to the dominant topic. The minimum value of this difference is 0, the maximum is 66, the mean is 14.96, and the median is 10. Similarly, I computed the difference in the weight of the dominant topic in the customer's messages and the weight of the same topic in the bot's messages. The minimum difference is 0, the maximum difference is 65, the mean is 14.67, and the median is 7.8.

I next aim to measure whether the similarity between the customer and the retailer representative—the agent, the bot, or both (in case of a chat that the bot transfers to an agent)—boosts *Customer engagement*. I measure *Customer engagement* as the average number of seconds the customer takes between consecutive messages. This measure complements prior research (e.g., Packard and Berger 2021) that has used customer satisfaction rating and purchase amount to measure the outcomes of customer-agent interactions.

Available Set of Features. I have a total of 330 features available to predict *Customer engagement*. I arrange these 330 features into three groups. First is 26 numeric customer-level and chat-specific features. Customer-level features include the unique customer ID maintained by the Salesforce Service Cloud database, the customer’s city, app or web browser, platform, browser language, screen resolution, mobile network, and the date and time of starting the chat. Chat-specific features are the number of seconds elapsed between the customer asking for a chat and the time the customer waits before being matched to a bot or an agent. These features also include whether the chat session involves a bot and the average number of seconds the operator (the bot, agent, or both) takes to respond. The second group includes the four “derived” numeric features for each of the up to three writers (i.e., customer and agent and/or bot) in a chat, totaling to 12 features. The four features for each writer are (a) the number of messages in the entire chat, (b) the number of words in the entire chat, (c) the type-token ratio (TTR), and (d) the number of questions. The third group comprises 300 text-based features created from vectorization (specifically, *Word2Vec* algorithm) of lemmatized customer messages, lemmatized bot messages, and lemmatized agent messages. In summary, my usable data set includes 292,473 rows and 338 features for my XGBoost.

Feature Selection. As the name suggests, *feature selection* reduces the number of features to conserve the computational cost of building a machine-learning model and simultaneously improves the model's performance (Chu et al. 2012).

I used a library named *featurewiz*, a fast feature selection library. The library proceeds in two steps. First, it identifies columns that have high information scores and the least correlation with one another. Second, it runs the extreme gradient boosting algorithm (XGBoost) on the identified columns to determine the top 143 features. I used these 143 features to build my XGBoost models.

XGBoost. I built three XGBoost models, which I present as Model 1 (M1), M2, and M3. Each model uses customer engagement as the response variable. M1 is my base model, which includes 143 features. I split it into a 75:25 ratio. That is, I used 75% of the observations to train my XGBoost and the remaining 25% to test it.

My Model 2 adds to Model 1 the indicator variable *Same Dominant Topic*, which equals 1 if the customer's dominant topic is the same as the agent's, the bot's, or either operator's (for chats where the bot transferred the customer to an agent). Thus, M2 trains on 144 features. Model 3 adds two continuous variables to Model 2. The first of these two variables is the difference between the weight of the dominant topic in the customer's messages and the weight of the same topic in the agent's messages. The second feature is the same difference measure, albeit between the customer and the bot. Thus, M3 trains on 146 features. Using the differences features, I test whether similarity in relative weight boosts customer engagement.

Results

I collected each model's two error metrics: root mean square error (RMSE) and mean absolute error (MAE). M1 reports an RMSE of 12.86. These low values indicate that the model

has a low error and is thus reasonably accurate. M2 improves by lowering the RMSE to 12.76.

Lastly, M3 reports an RMSE of 12.71. Overall, I conclude that using a difference measure for the dominant topic is as accurate as using an indicator measure.

DISCUSSION

As my check on the top 25 retailers in the United States revealed, the chat feature is popular. Text is unobtrusive and allows asynchronous conversation between two parties. In that sense, the text is a superior modality of expression to a face-to-face or a video/audio conversation that must be asynchronous (Berger, Rocklage, and Packard, 2022). As text becomes a more pervasive means for customers to interact with companies, more and more companies will introduce the chat feature on their websites, apps, and social media pages. Further, customers' demand for 24×7 and prompt service means that companies would use bots for first-line customer service. Should all these trends materialize, my research offers timely contributions to theory and practice.

Implications for Theory

My research extends the multidisciplinary evidence on chatbots. Prior evidence has documented customer response to chatbot features, such as humanization (Crolic et al. 2022; Schanke, Burtch, and Ray 2021), responsiveness and conversational tone (Jiang et al. 2022), warmth versus competence (Kull, Romero, and Monahan 2021) and emotions (Yun and Park 2021) in a bot's messages, and firm's disclosure that the customer is chatting with a bot (Luo et al. 2019; Mozafari, Weiger, and Hammerschmidt 2021). I contribute by showing that customers are open to using a bot for information about orders, coupons, and products, but ask for a transfer to a human agent for information on shipping and payment. In addition, I show that service agents are more similar (than the bot) to customers and that such similarity boosts customer

engagement. This finding asserts the role of employees in making bots effective (Gnewuch, Morana, Hinz, Kellner, and Maedche 2024) and extends prior research on linguistic style matching between conversation partners and the consequences of the match (Herhausen, Ludwig, Grewal, Wulf, and Schoegel 2019; Ireland and Pennebaker 2010; Ludwig, De Ruyter, Friedman, Brügger, Wetzels, and Pfann 2013; Moore and McFerran 2017).

Marketing is embracing two opportunities: customers' use of language in the marketplace and data science methods that can discover insights from this language (Packard et al. 2021). My research contributes to both opportunities. Specifically, in using chat transcripts as a novel source of marketing data, I take empirical marketing and consumer research to uncharted but consequential territory. This source of data is novel in multiple ways. First, unlike user-generated content on social media platforms, the chat transcripts are neither frivolous nor anonymous, indicating that one can obtain actionable insights from this text. Second, these chats characterize longitudinal and evolving problem-oriented conversations between a customer and a company representative. These messages are exchanged continuously over a period, offering an opportunity to study how the conversation evolves back and forth. Third, these chats involve both the customer and the marketer and have the customers' and the agent's satisfaction, time, and money on the line.

Implications for Managers

My three analyses offer three insights to managers. First, Analysis 1 reports that the bot wrote more words and a larger vocabulary than a human agent. This finding is counterintuitive because one would expect the bot to write fewer, and more precise words. Managers may use these findings to compare a bot's language with a human agent's and ensure that the bot's messages are efficient (i.e., use fewer words) and effective (i.e., use precise words). Further,

Analysis 1 discovers the key terms and topics from the corpus of customer messages, informing managers that customers are using the chat feature for six reasons. Managers may focus on the top three reasons and improve the feature's effectiveness in delivering value to the customer on these features. Such focused attention is better than improving customer service for all reasons. Second, Analysis 2 reports that agents are more similar to customers than bots. This finding makes a case for information technology companies—such as Salesforce, Microsoft, Google, and Amazon—to build more intelligence into their chatbots. One would expect that the advancements in large language models would allow a chatbot to dynamically learn from the customer's language use and mimic that use. Third, Analysis 3 demonstrates that a customer is more engaged if the operator's dominant topic matches the customer's (Packard et al. 2021). This intuitive and simple insight asserts the tangible outcomes of content-level matching.

Limitations and Future Research

My research suffers from at least three limitations, each of which merits future attention. First, I club all messages from a customer or the operator into one bag, discounting their temporal evolution. Future research may consider a longitudinal analysis, examining how a customer's first message shapes the agent's first message, how that impacts the customer's second message, and so on. Such a dynamic study of conversation would elicit nuances of how language impacts business outcomes. Second, I focus on language. Consumer psychologists have provided the theory (e.g., Luangrath, Peck, and Barger 2017) and the method (e.g., Luangrath, Xu, and Wang 2022) on how paralanguage can impact users and receivers of language. Future research can extend my research and examine how the two writers' use of paralanguage impacts, for example, customer engagement in a chat. Third, the lack of more detailed data prevented me from measuring the customer's and the agent's satisfaction rating for each customer-agent chat

or the customer's future transactions (Packard and Berger 2021). Organizational behaviorists and consumer psychologists may consider relating the dynamics of the conversation to not only the customer's satisfaction but also that of the agent. Fourth, customers can "chat" with a bot using voice. Psychologists have shown that speaking may have different impacts than writing (Berger, Rocklage, and Packard 2022). Future research may test whether chatting by voice offers different insights and impacts than chatting by text. Fifth, I use linguistic style matching to determine similarity. The short messages lack emotions and cognition, preventing me from exploring other linguistic features. Future research may measure similarity in these dimensions and relate it to customer engagement. Sixth, I sourced data from one retailer in a pre-large language model (pre-LLM) era. Future research may source data from alternative avenues and post-LLM period and inform academics and managers of whether the customer-chatbot conversation has evolved.

In summary, my research offers a novel contribution to research at the intersection of marketing, language, and computational methods while providing avenues that future research can build on.

REFERENCES

Berger, J., Kim, Y. D., & Meyer, R. (2021). What makes content engaging? How emotional dynamics shape success. *Journal of Consumer Research*, 48, 235–250.

Berger, J., Packard, G., Boghrati, R., Hsu, M., Humphreys, A., Luangrath, A., Moore, S., Nave G., Olivola C., & Rocklage, M. (2022). Wisdom from words: marketing insights from text. *Marketing Letters*, Forthcoming.

Berger, J., & Packard, G. (2023). Wisdom from words: The psychology of consumer language. *Consumer Psychology Review*, 6(1), 3-16.

Berger J, Rocklage MD, Packard G (2022) Expression modalities: How speaking versus writing shapes word of mouth. *Journal of Consumer Research*. Forthcoming.

Berger, J., Humphreys, A., Ludwig, S., Moe, W. W., Netzer, O., & Schweidel, D. A. (2020). Uniting the tribes: Using text for marketing insight. *Journal of Marketing*, 84(1), 1–25.

Carnevale, M., Luna, D., & Lerman, D. (2017). Brand linguistics: A theory-driven framework for the study of language in branding. *International Journal of Research in Marketing*, 34(2), 572-591.

Chu, C., Hsu, A.L., Chou, K.H., Bandettini, P., Lin, C. and Alzheimer's Disease Neuroimaging Initiative, 2012. Does feature selection improve classification accuracy? Impact of sample size and feature selection on classification using anatomical magnetic resonance images. *Neuroimage*, 60(1), 59-70.

Crolic, C., Thomaz, F., Hadi, R., & Stephen, A. T. (2022). Blame the bot: anthropomorphism and anger in customer–chatbot interactions. *Journal of Marketing*, 86(1), 132-148.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of Experimental Psychology: General*, 144(1), 114.

Fokina, M. (2023), 10+ crucial AI customer segment statistics (2024).
<https://www.tidio.com/blog/ai-customer-service-statistics/>

Golder, P. N., Dekimpe, M. G., An, J. T., van Heerde, H. J., Kim, D. S., & Alba, J. W. (2023). Learning from data: An empirics-first approach to relevant knowledge generation. *Journal of Marketing*, 87(3), 319-336.

Gnewuch, U., Morana, S., Hinz, O., Kellner, R., & Maedche, A. (2024). More than a bot? The impact of disclosing human involvement on customer interactions with hybrid service agents. *Information Systems Research*, 35(3), 936-955.

Grewal, D., Hulland, J., Kopalle, P. K., & Karahanna, E. (2020). The future of technology and marketing: A multidisciplinary perspective. *Journal of the Academy of Marketing Science*, 48, 1-8.

Herhausen, D., Ludwig, S., Grewal, D., Wulf, J., & Schoegel, M. (2019). Detecting, preventing, and mitigating online firestorms in brand communities. *Journal of marketing*, 83(3), 1-21.

- Hoffman, D. L., Moreau, C. P., Stremersch, S., & Wedel, M. (2022). The rise of new technologies in marketing: A framework and outlook. *Journal of Marketing*, 86(1), 1-6.
- Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research*, 21(2), 155-172.
- Huang, M. H., & Rust, R. T. (2024). The caring machine: Feeling AI for customer care. *Journal of Marketing*, 88(5), 1-23.
- Humphreys, A., & Wang, R. J. H. (2018). Automated text analysis for consumer research. *Journal of Consumer Research*, 44(6), 1274-1306.
- Ireland, M. E., & Pennebaker, J. W. (2010). Language style matching in writing: synchrony in essays, correspondence, and poetry. *Journal of Personality and Social Psychology*, 99(3), 549.
- Jiang, H., Cheng, Y., Yang, J., & Gao, S. (2022). AI-powered chatbot communication with customers: Dialogic interactions, satisfaction, engagement, and customer behavior. *Computers in Human Behavior*, 107329, 134 (September), <https://doi.org/10.1016/j.chb.2022.107329>.
- Jussupow, E., Benbasat, I., & Heinzl, A. (2020). Why are we averse towards algorithms? A comprehensive literature review on algorithm aversion. *Twenty-Eighth European Conference on Information Systems (ECIS2020) – A Virtual AIS Conference*.
- Luangrath, A. W., Peck, J., & Barger, V. A. (2017). Textual paralanguage and its implications for marketing communications. *Journal of Consumer Psychology*, 27(1), 98-107.
- Luangrath A. W., Xu Y., Wang T. (2022) Paralanguage Classifier (PARA): An algorithm for automatic coding of paralinguistic nonverbal parts of speech in text. Working Paper.
- Ludwig, S., De Ruyter, K., Friedman, M., Brüggem, E. C., Wetzels, M., & Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. *Journal of Marketing*, 77(1), 87-103.
- Kronrod, A. (2022). Language research in marketing. *Foundations and Trends® in Marketing*, 16(3), 308-421.
- Kull, A. J., Romero, M., & Monahan, L. (2021). How may I help you? Driving brand engagement through the warmth of an initial chatbot message. *Journal of Business Research*, 135, 840-850.
- Luo, X., Tong, S., Fang, Z., & Qu, Z. (2019). Frontiers: Machines vs. humans: The impact of artificial intelligence chatbot disclosure on customer purchases. *Marketing Science*, 38(6), 937-947.

Mariani, M. M., Hashemi, N., & Wirtz, J. (2023). Artificial intelligence empowered conversational agents: A systematic literature review and research agenda. *Journal of Business Research*, 161, 113838.

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

Mozafari, N., Weiger, W. H., & Hammerschmidt, M. (2021). Trust me, I'm a bot—repercussions of chatbot disclosure in different service frontline settings. *Journal of Service Management*.

Niederhoffer, K. G., & Pennebaker, J. W. (2002). Linguistic style matching in social interaction. *Journal of Language and Social Psychology*, 21(4), 337-360.

Packard, G., & Berger, J. (2021). How concrete language shapes customer satisfaction. *Journal of Consumer Research*, 47(5), 787-806.

Packard, G., & Berger, J. (2024). The emergence and evolution of consumer language research. *Journal of Consumer Research*, 51(1), 42-51.

Packard, G., Berger, J., Boghrati, R., Yeomans, M., Minson, J., Collins, H., Gino, F., Donnelly, G., Hurst, K., Sintov N., and Li, Y. (2021). Understanding Consumer Conversations. *ACR North American Advances*, 49, eds. Tonya Williams Bradford, Anat Keinan, and Matthew Matthew Thomson, Duluth, MN: Association for Consumer Research, Pages: 780-785.

Packard, G., Moore, S. G., & McFerran, B. (2018). (I'm) happy to help (you): The impact of personal pronoun use in customer–firm interactions. *Journal of Marketing Research*, 55(4), 541-555.

Pentina, I., Xie, T., Hancock, T., & Bailey, A. (2023). Consumer–machine relationships in the age of artificial intelligence: Systematic literature review and research directions. *Psychology & Marketing*, 40(8), 1593-1614.

Pogacar, R., Lowrey, T. M., & Shrum, L. J. (2017). The influence of marketing language on brand attitudes and choice. In *The Routledge companion to consumer behavior* (pp. 263-275). Routledge.

Pogacar, R., Shrum, L. J., & Lowrey, T. M. (2018). The effects of linguistic devices on consumer information processing and persuasion: A language complexity× processing mode framework. *Journal of Consumer Psychology*, 28(4), 689-711.

Schanke, S., Burtch, G., & Ray, G. (2021). Estimating the impact of “humanizing” customer service chatbots. *Information Systems Research*, 32(3), 736-751.

Servion Global Solutions (2020), “AI Will Power 95% of Customer Interactions by 2025,” <https://www.financedigest.com/ai-will-power-95-of-customer-interactions-by-2025.html>.

Shawar, B. A., & Atwell, E. S. (2005). Using corpora in machine-learning chatbot systems. *International Journal of Corpus Linguistics*, 10(4), 489-516.

Yun, J., & Park, J. (2022). The effects of chatbot service recovery with emotion words on customer satisfaction, repurchase intention, and positive word-of-mouth. *Frontiers in Psychology*, 13, 922503.

Woolley, K., & Fishbach, A. (2017). A recipe for friendship: Similar food consumption promotes trust and cooperation. *Journal of Consumer Psychology*, 27(1), 1-10.

Insights from Customers' Chats with Bots and Human Agents

Web Appendix A

Customer-Agent Transcript

Raw Customer-Agent Transcript

<p align="center">Chat Started: Friday, September 18, 2020, 12:37:25 (-0700)</p><p align="center">Chat Origin: Live Agent Chat</p><p align="center">Agent Custom Agent Name</p>(29m 3s) Custom Agent Name: Custom Agent Greetng
(29m 18s) Custom Agent Name: how can i help you?
(29m 49s) Mary: Hello I'm wanting to verify my order because i got an email it was delivered and I don't see it
(30m 5s) Mary: Order #0008187449271
(30m 18s) Custom Agent Name: i'll be happy to assist you with verifying your order, 1 moment while i look up your order
(30m 45s) Mary: Ok great thank you
(32m 23s) Custom Agent Name: here is your tracking number 61290985582820143413, it says on FedEx that Left at front door. Signature Service not requested.
(33m 4s) Custom Agent Name: you can contact FedEx directly & see if they can provide you with more information
(34m 42s) Mary: Ok I was trying to locate the tracking number online and didn't let me. Ok so contact fedex and then what ? Explain the situation. What happens next just wait to hear from fedex
(36m 12s) Custom Agent Name: you would have to see what FedEx tells you & any time frame they provide to you
(37m 0s) Custom Agent Name: if after contacting them & they are still unable to locate your order, please contact us back & we will be happy to assist you with a refund
(37m 33s) Mary: Ok sounds good I will contact fedex right now thank you
(37m 54s) Custom Agent Name: your welcome, thank you for shopping, have great day

Table A1A: Customer-Agent Transcript, Separated into Messages

| Writer | Message |
|----------|--|
| Agent | how can i help you? |
| Customer | Hello I'm wanting to verify my order because i got an email it was delivered and I don't see it. |
| Customer | #0008187449271 |
| Agent | I'll be happy to assist you with verifying your order, 1 moment while i look up your order |
| Customer | Ok great thank you |
| Agent | here is your tracking number 61290985582820143413, it says on FedEx that Left at front door. Signature Service not requested. |
| Agent | you can contact FedEx directly & see if they can provide you with more information |
| Customer | Ok I was trying to locate the tracking number online and didn't let me. Ok so contact fedex and then what? Explain the situation. What happens next just wait to hear from fedex |
| Agent | you would have to see what FedEx tells you & any time frame they provide to you |
| Agent | if after contacting them & they are still unable to locate your order, please contact us back & we will be happy to assist you with a refund |

| | |
|----------|---|
| Customer | Ok sounds good I will contact fedex right now thank you |
| Agent | our welcome, thank you for shopping, have great day |

Customer-Bot Transcript

Raw Customer-Bot Transcript

<p align="center">Chat Started: Tuesday, December 15, 2020, 11:18:42 (+0000)</p>(1s)
 Customer Service: Welcome to Shoe Carnival customer service.
(3s) Customer Service: I am looking forward to helping you today.
(4s) Customer Service: {ChatWindowButton:Check my order status,Buy Online Pickup in Store,Check Gift Card Balance,Redeem a Gift Card,Log in to Shoe Perks,Unsubscribe to Email,Other questions}
(7s) Visitor: Check my order status
(9s) Customer Service: Sure, let's check on your order.
(10s) Customer Service: Do you have an order number?
(12s) Customer Service: {ChatWindowButton:Yes,No}
(14s) Visitor: No
(16s) Customer Service: Let me connect you to someone that can help
(18s) Customer Service: I am sorry there's nothing I can help you with.

Let me connect you with another customer agent.
<p align="center">Agent Chatbot failed to transfer the chat to button Chat Bot Group</p><p align="center">Agent Chatbot chat resumed with BotId 0Xx3h000000kHhT</p>(20s)
 Customer Service: Unfortunately, there are no agents available at the moment. Please try again during our business hours.
(1m 1s)

Table A1B: Customer-Bot Transcript, Separated into Messages

| Writer | Message |
|----------|--|
| Bot | Welcome to Shoe Carnival customer service. |
| Bot | I am looking forward to helping you today. |
| Bot | Options: Check my order status,Buy Online Pickup in Store,Check Gift Card Balance,Redeem a Gift Card,Log in to Shoe Perks,Unsubscribe to Email,Other questions |
| Customer | Check my order status |
| Bot | Sure, let's check on your order. |
| Bot | Do you have an order number? |
| Bot | Options: Yes, No |
| Customer | No |
| Bot | Let me connect you to someone that can help |
| Bot | I am sorry there's nothing I can help you with. Let me connect you with another customer agent. Agent Chatbot failed to transfer the chat. Chatbot chat resumed. |
| Bot | Unfortunately, there are no agents available at the moment. Please try again during our business hours. |

Customer-Bot-Agent Transcript

Raw Transcript

<p align="center">Chat Started: Tuesday, December 15, 2020, 07:03:42 (-0600)</p>(2s)
 Customer Service: Welcome to Shoe Carnival customer service.
(4s) Customer Service: I am looking forward to helping you today.
(5s) Customer Service: {ChatWindowButton:Check my order status,Buy Online Pickup in Store,Check Gift Card

Balance, Redeem a Gift Card, Log in to Shoe Perks, Unsubscribe to Email, Other questions}
(7s) Visitor: Check my order status
(9s) Customer Service: Sure, let's check on your order.
(10s) Customer Service: Do you have an order number?
(11s) Customer Service: {ChatWindowButton: Yes, No}
(15s) Visitor: Yes
(18s) Customer Service: What is your order number?
(21s) Visitor: 88436104
(24s) Customer Service: There seems to be a problem with that number. Let's try that again.
(25s) Customer Service: What is your order number?
(32s) Visitor: 88436104
(35s) Customer Service: Let me connect you to some one that help
(37s) Customer Service: I am sorry there's nothing I can help you with.

Let me connect you with another customer agent.
<p align="center">Agent Chatbot successfully transferred the chat to button Chat Bot Group</p><p align="center">Chat Started: Tuesday, December 15, 2020, 07:04:26 (-0600)</p><p align="center">Chat Origin: Chat Bot Group</p><p align="center">Chat Transferred From Customer Service To Kasey</p>(44s) Kasey: Hello, how can I assist you today?
(3m 20s) Visitor: i still haven't gotten my order
(3m 29s) Visitor: 88436104
(4m 47s) Visitor: are you there
(5m 21s) Kasey: Yes I am! I am checking on this now!
(7m 26s) Kasey: I apologize for the delay with your order. It looks like the order got stuck by an error when processing through our system. I am going to escalate this immediately to the proper department so that the order can get pushed through.
(7m 53s) Visitor: ok, do you think it will make it for Christmas
(8m 27s) Kasey: I am asking that it be expedited on us. It should make it!
(9m 3s) Visitor: ok thank you for your help
(9m 40s) Kasey: You're welcome! Is there anything else I can assist you with today?
(9m 47s) Visitor: no thank you
(1m 1s)

Table A1C: Customer-Bot-Agent Transcript, Separated into Messages

| Writer | Message |
|--|---|
| Bot | Welcome to Shoe Carnival customer service. |
| Bot | I am looking forward to helping you today. |
| Customer | Check my order status |
| Bot | Sure, let's check on your order. |
| Bot | Do you have an order number? Yes, No |
| Customer | Yes |
| Bot | What is your order number? |
| Customer | 88436104 |
| Bot | There seems to be a problem with that number. Let's try that again. |
| Bot | What is your order number? |
| Customer | 88436104 |
| Bot | Let me connect you to some one that help |
| Bot | I am sorry there's nothing I can help you with. Let me connect you with another customer agent. |
| Chatbot successfully transferred the chat | |
| Agent | Hello, how can I assist you today? |
| Customer | i still haven't gotten my order |
| Customer | 88436104 |
| Customer | are you there |
| Agent | Yes I am! I am checking on this now! |

| | |
|----------|---|
| Agent | I apologize for the delay with your order. It looks like the order got stuck by an error when processing through our system. I am going to escalate this immediately to the proper department so that the order can get pushed through. |
| Customer | ok, do you think it will make it for Christmas |
| Agent | I am asking that it be expedited on us. It should make it! |
| Customer | ok thank you for your help |
| Agent | You're welcome! Is there anything else I can assist you with today? |
| Customer | no thank you |

Preprocessing

Table A2A: Customer Messages: Raw, Clean, Lemmatized

| Customer Messages – Raw | Customer Messages – Clean | Customer Messages – Lemmatized |
|---|--|---|
| Good Morning, my order 88679177 was canceled, when will I be reimbursed for that purchase? (3m 56s) It was not out of stock and I reordered the item. (8m 38s) Thanks (9m 46s) | good morning order canceled reimbursed purchase stock reordered item | good morning order cancel reimburse purchase stock reorder item |
| | | |

Table A2B: Agent Messages: Raw, Clean, Lemmatized

| Agent Messages – Raw | Agent Messages – Clean | Agent Messages – Lemmatized |
|---|--|--|
| Hello, how can I assist you today? (2m 38s) I can check this for you (6m 21s) It looks like the item you ordered was out of stock when you ordered it. Unfortunately it is still out of stock. Your order was canceled yesterday. Refunds are processed in 3-5 days (7m 39s) I apologize for this, our system seems to be presenting false information. However, you will be refunded in 3-5 days (8m 51s) Glad I could help. Have an amazing day :) (1m 1s) | hello assist today check looks like item ordered stock ordered unfortunately still stock order canceled yesterday refunds processed days apologize system seems presenting false information however refunded days glad could help amazing day | hello assist today check look like item order stock order unfortunately still stock order cancel yesterday refund process day apologize system seem present false information however refunded day glad could help amazing day |

Table A2C: Bot Messages: Raw, Clean, Lemmatized

| Bot Messages – Raw | Bot Messages – Clean | Bot Messages – Lemmatized |
|---|--|---|
| <p>Tuesday, December 15, 2020, 07:04:06 (-0600)</p>(2s) Welcome to Shoe Carnival customer service. (3s) I am looking forward to helping you today. (4s) {ChatWindowButton:Check my order status,Buy Online Pickup in Store,Check Gift Card Balance,Redeem a Gift Card,Log in to Shoe Perks,Unsubscribe to Email,Other questions} (18s) I am sorry there's nothing I can help you with. Let me connect you with another customer agent. <p align="center">Agent Chatbot successfully transferred the chat to button Chat Bot Group</p><p align="center">Chat Started: Tuesday, December 15, 2020, 07:05:07 (-0600)</p><p align="center">Chat Origin: Chat Bot Group</p><p align="center">Chat Transferred From Customer Service To Agent</p>(1m 1s)</p> | <p>welcome customer service looking forward helping today check order status buy online pickup store check gift card balance redeem gift card log shoe perk unsubscribe email question sorry nothing help connect another customer agent agent chatbot successfully transfer chat started chat bot chat transferred customer service agent</p> | <p>welcome customer service look forward help today check order status buy online pickup store check gift card balance redeem gift card log shoe perk unsubscribe email question sorry nothing help connect another customer agent agent chatbot successfully transfer chat start chat bot chat transfer customer service agent</p> |