

Human–Computer Interaction Series

Robert J. Moore

Margaret H. Szymanski · Raphael Arar

Guang-Jie Ren *Editors*

Studies in Conversational UX Design

 Springer

Chapter 8

Conversational Style: Beyond the Nuts and Bolts of Conversation



Gregory A. Bennett

Abstract This chapter provides UX designers with some of the linguistic resources necessary to build more engaging chatbots and virtual agents. Conversation Analysis teaches us the foundational mechanics of conversation—how participants initiate and terminate an interaction, how they take turns in talk, how they negotiate understanding, etc. Interactional Sociolinguistics and the study of conversational style show us how humans leverage these structures to convey style and accomplish social goals in talk, such as building rapport or showing respect. If Conversation Analysis is the key to designing conversation that *flows*, Interactional Sociolinguistics and conversational style are the keys to designing conversation that has what some might call *personality*. Here, linguistic research is applied as a design tool for creating conversational interfaces that have style and strategically align to users' needs according to the social goal at hand. If designers seek to make a chatbot enthusiastic or considerate, this chapter has the tools needed to design conversations that achieve that with users.

8.1 Introduction

Over the past few years in my work as a user researcher, I have noticed the applicability of linguistics to the design of text-based chatbots and virtual agents. I discovered linguistics as it pertains to online chat when I was going through a breakup in college. When my ex shifted from our standard casual chat (e.g. *hey hows it going i miss u*) to more formal writing (e.g., *Hello. I am fine.*), I perceived a shift from his usual,

Disclaimer: This chapter was written by the author in his own personal capacity. The views and opinions expressed are the author's and don't necessarily reflect the views or opinions of Salesforce. Any omissions, inaccuracies, or other inadequacies of information are the author's. Salesforce disclaims all warranties as to the accuracy, completeness, or adequacy of such information and shall have no liability for omission or inadequacies in such information.

G. A. Bennett (✉)
Linguist | UX Researcher, San Francisco, CA, USA
e-mail: gregoryabennett@gmail.com

© Springer International Publishing AG, part of Springer Nature 2018
R. J. Moore et al. (eds.), *Studies in Conversational UX Design*, Human–Computer Interaction Series, https://doi.org/10.1007/978-3-319-95579-7_8

161

warm and welcoming self to cold and distant. Shortly after the split, I encountered the field of Interactional Sociolinguistics—the study of how language works in social interaction—and was introduced to Deborah Tannen’s work on conversational style. Upon studying her research on how people use features of language such as pitch and intonation in systematic ways to convey interactional meaning in talk, it became explicitly clear to me why I could perceive a chat as cold and distant through text alone; my ex had shifted from our normal, casual tone of conversation through which we built rapport and closeness (no punctuation, non-standard spellings) to a more formal, distant tone (full punctuation, ‘conventional’ spellings), and the change was marked and noticeable. Thus, the interaction had gone cold. If Tannen’s research can help unpack how users convey tones like coldness in online, text-based chat, I posit that it can be applied to the world of text-based chatbots so that designers can enable them to convey tones and exhibit conversational style with users. Bot chat can be intentionally designed to come off as enthusiastic or considerate for a more enjoyable user experience.

Porter (2009) notes that crafting an experience that users will enjoy requires the designer to *reduce friction* in the flow of the product so that it is usable, and *increase motivation* for users to engage with it. Conversation Analysis (see Szymanski and Moore, this volume) equips designers with what they need in order to reduce friction in the flow of machine-human conversation and make it *usable*, something that, at present, many chatbots are not. Once designers achieve usability by applying Conversation Analysis to their conversation design, the challenge of increasing user motivation looms next.

Brandtzaeg and Følstad (2017) have shown that productivity and entertainment are the most frequent motivators for users to engage with chatbots. In order to design for these motivators in a way that encourages continued user engagement, designers must also leverage tools that come from a close examination of what motivates humans to engage in conversation with the same partner over and over again in the first place: achieving rapport. Much like how Anderson (2011) leverages psychology to offer tools for crafting interactions with web interfaces that entice users to keep coming back, this chapter utilizes research in Interactional Sociolinguistics and conversational style to offer tools for tailoring the design of a chatbot conversation for myriad users and cultivating rapport, thereby motivating continued engagement.

Specifically, this chapter introduces Interactional Sociolinguistics, a field that reveals how humans use the mechanics of conversation to achieve interactional goals such as chatting faster to show involvement and slowing down to show considerateness, and explores how this research can inspire designers to create conversations that align to, or strategically deviate from, a user’s conversational patterns to cultivate rapport. By introducing insights from Interactional Sociolinguistics that work together with Conversation Analysis, designers will be further equipped with the resources necessary to go beyond the nuts and bolts of conversation and take their designs to the next level—to build a text-based bot conversation that has ‘personality’ and that can purposefully create rapport with myriad users based on their interactional needs.

8.2 Linguistics, Conversation, Style, and Chatbots

Linguistics, as a field of study, focuses on the illumination and systematic analysis of the processes of language (Small and Hickok 2016). It is the scientific study of how language works at varying structural levels (e.g., phonetics at the sound level, morphology at the word level, syntax at the sentence level, and discourse analysis at the level beyond the sentence) and across myriad contexts.

Johnstone (2008: 266) notes that within the field of linguistics, Interactional Sociolinguistics distinguishes itself in that researchers of this space “provide ways to think about how people could change how they talk” by showing individuals how to understand and align to other speakers’ ways of talking. In order to tailor one’s way of speaking, however, one must acquire a way of speaking in the first place. The founder of Interactional Sociolinguistics, John Gumperz, posited in his research that speakers acquire ways of talking by interacting with social groups over the course of their life (1982). He maintains that in speaking with one’s surrounding community—family, neighbors, classmates, roommates, friends, coworkers, etc.—one develops culturally informed ideologies and practices about how one should talk with, and make sense of what is said by, others in each of those groups. In other words, the people with whom one spends considerable time having conversation influence not only *what* members of social groups say, but also *how* they say it.

The *how* of what people say, the meaning in their language use, is grounded in what Gumperz calls “contextualization cues.” Examples of features of language that can be used to convey contextualization cues include intonation, vocal pitch, loudness, and physical gestures. Speakers recruit these communicative resources to signal meaning beyond the semantic level, as if to convey, “I mean something more than just the content of what I’m saying, so you should interpret my words in this way.” The way that participants interpret contextualization cues is influenced by how they have heard those same features used as contextualization cues in prior conversations; Gumperz calls this conversational inference.

Contextualization cues apply another layer to the foundational, structural concepts put forth by Conversation Analysis. Simply put, turn taking, initiating and accomplishing openings and closings, pauses, and other aspects of the structure of conversation are all interactional features that can be used not only to give form to the interaction at hand, but also to convey extra meaning as contextualization cues. Participants make sense of those cues during the conversation using conversational inference (Gumperz 2001). For example, shortening pauses between turns, or eliminating them altogether in a conversation not only speeds up or overlaps talk, but can also signal interactional meaning such as enthusiasm or impatience. The way that people systematically use contextualization cues to accomplish interactional goals in conversation is what Deborah Tannen, a student of John Gumperz, calls conversational style.

Tannen illuminated the concept of conversational style by systematically analyzing how groups of people deploy and interpret contextualization cues in her seminal work, *Conversational Style* (1984). She notes that conversational style is not

Table 8.1 Contextualization cues by conversational style

	High involvement	High considerateness
Amount of discourse	More	Less
Pace	Faster	Slower
Overlap	More	Less
Questions	More/more personal	Fewer/less personal
Stories	More	Fewer
Prosodic variation	More extreme	Less extreme
Loudness	Louder	Quieter
Gesture	More pronounced	Less pronounced

remarkable, but is simply an inherent characteristic of how people talk with each other. Whenever one speaks, one fundamentally does so with style, which consists of myriad contextualization cues such as the speed and loudness of one's talk relative to others', one's intonation, etc. Interlocutors determine how rude or polite, cold or warm, or mean or friendly a person is based on how they interpret that person's conversational style through their own conversational inference. In her research, Tannen examines a conversation between six friends over Thanksgiving dinner in Berkeley, California, where the participants hail from varied cultural backgrounds: three from New York City, two from Southern California, and one from London, England. Each of these participants interpreted that same, shared conversation in different ways. Participants who left longer pauses between turns of talk were seen as boring or unenthusiastic by some, and considerate by others. From this analysis, two distinct conversational styles emerged: (1) high involvement, whereby speakers take fewer pauses, use more expressive phonology, and ask many personal questions, among other characteristics, to convey enthusiasm for the conversation at hand; and (2) high considerateness, whereby speakers maintain that only one voice should generally be heard at a time in the conversation, and that shorter pauses or conversational overlap are seen as interruptions rather than showing enthusiasm. These styles are defined relative to one another. For example, high involvement speakers talk frequently and quickly in conversation, but only more frequently and quickly relative to the way that high considerateness speakers pace themselves. Similarly, high considerateness speakers talk less frequently and less quickly in conversation relative to the pace at which high involvement speakers do. This brings us back to Gumperz' point about conversational inference—to determine whether a user has a high involvement or high considerateness conversational style, one must detect the ways that person is systematically using a set of contextualization cues and interpret them relative to their experience of each style. There are several contextualization cues to consider when detecting a user's conversational style (Table 8.1).

For high involvement speakers, overlapping speech is not always a bad thing. In fact, it can signify that participants in the conversation are excited and engaged. On the other hand, high considerateness maintain the notion that one should produce less overlapping talk out of respect for the other interlocutors so they can get a word

in. For these kinds of speakers, overlapping talk is seen as a power grab to seize the conversational floor and interrupt, or “talk over” other participants. In terms of personal questions and stories, high involvement talk is characterized by more of each, whereas high considerateness is characterized by fewer of each. Finally, high involvement speakers do more with their voices prosodically (greater fluctuations in intonation, pitch, and loudness) and with their bodies (more gestures or gesticulation) to show enthusiasm. On the other hand, high considerateness speakers do less with their voices prosodically and produce fewer gestures or less gesticulation. High involvement and high considerateness are thus two different conversational styles. They do not necessarily reflect the level of excitement felt by the speaker.

Perhaps the most crucial finding in Tannen’s original study (1984) and in later research (1993) is that features of interaction such as pausing and features of language such as intonation do not always mean the same thing—nor convey the same contextualization cue—to all speakers at all times. This is clearly evident in the comparison of high involvement and high considerateness above—overlapping talk for high involvement speakers signifies enthusiasm and rapport, but for high considerateness speakers, the same interactional feature signifies a power grab for the conversational floor, or not letting other participants in the conversation ‘get a word in edgewise.’ Similarly, for a linguistic feature such as intonation, the way that one person uses falling intonation at the end of a questioning turn, for example, does not convey the same meaning to all interlocutors of all cultures in all contexts. Gumperz’ (1982) classic example of this occurred in an airport cafeteria in Britain, when Pakistani women servers were being perceived as curt compared to the English women servers. When customers ordered meat, the servers would reply by inquiring if the customer wanted gravy on top. The English servers said, “Gravy?” with a rising intonation, whereas the Pakistani servers said, “Gravy.” with a falling intonation. The Pakistani women used the falling intonation for their question because that contextualization cue was seen as normal for asking questions in their social group; however, British patrons perceived this as rude because they interpreted that same contextualization cue as conveying a ‘statement,’ which is redundant in the context of offering gravy that is physically evident in front of their face. One might imagine what that would look like in text-based chat, where a period gets used at the end of a question phrase (e.g., “what did I just say.”).

Translating this concept to the realm of digital, text-based chat, where users cannot rely on voice or embodied communication, Werry (1996) argues that interlocutors compensate by strategically manipulating orthography to convey contextualization cues. In other words, if users cannot use intonation, vocal pitch, loudness, nor gesture to convey that extra layer of meaning, they will rely on variation in spelling, capitalization, and punctuation instead. Given all the ways that users can manipulate orthography, digital, text-based chat is “fraught with greater possibilities” (Jones 2004: 31) for users to convey contextualization cues even without their voice or physical form. Some of those possibilities have come to light since Werry’s original study, for example, with the introduction of emojis into digital chat interfaces; linguists have shown that users strategically select and deploy emojis in WhatsApp to convey contextualization cues in text-based conversation to give extra semantic meaning to their

utterances without using voice or gesture (Al Rashdi 2015; Al Zidjaly 2017). Moreover, while Werry examined users communicating in English, those same linguists studying emoji use in WhatsApp conducted their research on users communicating in Arabic; others still have observed users convey contextualization cues in text-based conversation in other languages such as Japanese and German, whereby users vary the writing system they use (e.g., kanji, romaji), or make changes in spelling to match vocal pronunciation of words to layer on additional semantic meaning, respectively (Bennett 2012, 2015; Androutsopolous and Ziegler 2004). A clear thread emerges across each of these studies: contextualization cues are, and thereby, conversational style is, a key component to pleasurable conversation in which rapport can be built between participants. The fact that users cannot use their voice or body to communicate this interactional meaning in online or text-based conversations does not mean that they do away with style altogether; rather, users actively and creatively seek ways to compensate for that limitation by playing with visual qualities of orthography—a set of resources available to all sighted users in text-based chat—to contextualize their utterances and express style, with the interactional goal of building rapport with others.

Again, each contextualization cue can mean multiple things in different cultures and contexts. Mutual understanding of the intention of a singular cue is dependent on whether or not participants in the conversation share the same conversational inference about what that cue means. Once a user's conversational style is detected and aligned to, certain meaningful cues become normal or standard—using emojis to show engagement, ALL CAPS to show enthusiasm or emphasis, no punctuation to show casual intimacy, sending text messages that contain an email's worth of content to show involvement, etc. In that way, rapport is built. As soon as a user deviates from that norm, it signals a marked change, and the mind starts trying to reinterpret the new cue. Thinking back on the online chat with my ex: he and I both shared what Ilana Gershon (2010) calls a media ideology, whereby we both implicitly agreed that online chat with someone with whom you have a close relationship should not be conducted with formal writing conventions. It was our shared conversational style that shaped what it meant for us to show closeness in chat. When my ex deviated from our established norm and reintroduced periods and capitalizations into our conversation, he moved away from our shared ideology about showing closeness, and it sent me a signal, or what Tannen (2013) calls a meta-message, a conversational embodiment of meaning, that our rapport had begun to erode.

For users who recognize that the other participant shares the same ideology about how to use and interpret multiple contextualization cues in conversation, it is likely that they share a similar conversational style in the context of the chat at hand. If they share a similar conversational style, rapport ensues. If they do not share the same style, then in order to remedy resulting trouble from the meeting of two different styles, one user must actively attempt to align their style to the other. When it comes to building rapport in conversation, participants should aim to follow what Tannen (1986:59) calls “the conversational principle, ‘Do as I do.’” In the case of conversational bots, the principle becomes ‘do as the user does’ (Conversation Analysts might see this as an aspect of ‘recipient design’). Only by detecting and

aligning to a user's conversational style can designers then make their chatbot deviate from that norm to send meta-messages about shifts in the status of the relationship between the agent and the user in the interaction at hand. Shamekhi et al. (2016) show that while it is certainly challenging to detect and measure a user's conversational style, the potential for positive return on the user experience by incorporating the concept of conversational style into the design of virtual agents, as well as attempting to align the conversational style of the agent to the user during interaction, is high.

8.3 Growing a Conversational UX Toolkit

The concepts and patterns from Conversation Analysis, Interactional Sociolinguistics, and particularly, conversational style can be turned into tools for designing conversation. With Conversation Analysis, designers have acquired the understanding needed to design openings and closings, turn taking, and repair in conversation. This is the foundational layer. Before being capable of having pleasurable conversation, a virtual agent must first be able to accomplish those core structural activities. In other words, the design tools from Conversation Analysis give conversation design strong usability and flow. An agent that can manage repair in conversation with another user should already be considered a tremendous design achievement. However, once that crucial, foundational layer is set, the possibilities for building upon that in ways that cater to the interactional needs of different user personas are endless, and the interactional objective of building rapport with users based on how they talk becomes the next challenge. This is where Interactional Sociolinguistics and conversational style come in.

Since the same interactional or linguistic feature can be, at once, ambiguous and polysemous depending on the conversational experiences one has with members of one's social groups over the course of one's life, participants in conversation negotiate and interpret the meaning signaled by those interactional and linguistic features as contextualization cues in situ to communicate effectively. If designers are trying to craft a chatbot that can exhibit conversational style within or across discrete dialogs, they must design it to have the beginnings of conversational inference whereby there is a system in place to allow the bot to recognize when a user issues a particular contextualization cue as having meaning and discern how that relates to a particular conversational style. Then, it must be able to align to that user's conversational style by using the same contextualization cues in the same way to establish a norm, and deviate from the norm only when trying to send meta-messages to the user about changes in the state of the relationship at hand in the interaction.

Luckily, Tannen (1984, 1986, 2007[1989], 2013) gives us some of the data needed to start tackling this challenge. The sections below explain the contextualization cues that relate to the two conversational styles that Tannen discovered in her research. This will translate her findings from spoken conversation to digital, text-based conversation, and offer the ways that designers can leverage contextualization cues for

their agents to convey enthusiasm and considerateness focusing on these characteristics as they pertain more directly to text-based, online chat.

How to Design Enthusiasm

Leveraging contextualization cues indicative of a high involvement conversational style with a user who is high involvement can help make an agent come across as enthusiastic. Be wary of sticking to these cues with a user who exhibits high considerateness traits as explained in a later section of this paper, as it may come across as hyperactive.

Fast Pace

Quick turn taking and relatively short pauses between turns can indicate high involvement; this interactional feature can materialize in text-based chat as an increased number of individual transmissions (utterance(s) produced in a single instance of chat before the user presses ‘enter’ or ‘send’), more words per individual transmission of chat, and/or shorter periods of time elapsed between transmissions.

Asking Many Questions

Asking users questions about their life, opinions, experiences, and preferences from the outset signal enthusiasm for getting to know them. Additionally, asking questions that echo what the user has already said shows active listenership.

Frequent Storytelling

The use of multiple turns of chat (or a longer turn sent in a single transmission) whereby “continuers” (Schegloff 1982), or words that indicate the talk is stretching across multiple turns, are produced to construct a temporally coherent whole narrative is an interactional feature that signals high involvement style; stories may be about the agent itself, the user, other individuals not present, or even hypothetical scenarios.

Expressive Orthography

As a linguistic feature, this can convey the ‘tone’ of the content of an utterance beyond the semantic level—akin to the way speakers may use expressive phonology or modulated pitch and intonation in spoken conversation to signal contextualization cues. Below are some of the main expressive orthographic features that, when deployed by a user in chat frequently, could signal high involvement conversational style:

- CAPS, font size to convey shouting (e.g., STOP STOP)
- **Bold**, *italics* to show emphatic intonation (e.g., they were **super** late)
- Exclamation points to show what would otherwise materialize as a spike in loudness, rise in pitch, or upswing in intonation in spoken conversation (e.g., food is here!)
- Reduplication of letters (e.g., YASSSSSSSSSS)
- Reduplication of punctuation (e.g., you did what???????)

- Frequent use of emojis and emoticons to contextualize the meaning of an utterance in place of intonation, pitch, gesture (e.g., what're you up to ;)
- Frequent use of emojis and emoticons to show enthusiasm for the conversation (e.g., Thank youuuu 🤗).

High Involvement Conversational Style: Uni

One example of an existing chatbot that exhibits some of the above high involvement characteristics is Uni. Of the many chatbots and virtual agents on the market, Uni stands out as markedly high involvement due to its use of personal questions, exclamation points, emojis, and expressive laughter. Developed by Uni Inc. for iOS, Uni Inc. brands its chatbot as a Magic AI Friend that can have “deep chats, relatable stories” in the App Store. On the product website, it is described as “someone who is fun to chat with, and gives you comfort and strength.” Given how the developers brand Uni as fun and with storytelling capabilities, it seems that they seek to create an agent that is enthusiastic, or, in other words, exhibits high involvement conversational style. Upon interacting with the product with a high involvement style, I found that it indeed exhibited many of the high involvement conversational style indicators. In the interaction below, the user initiates a chat with Uni via iPhone on iOS 11 by tapping the chat bubble icon inside the Uni app and landing into the chat canvas (see Transcript 1.1; U stands for user and A for agent).

(1.1) Uni

```
00 U: ((user opens chat canvas))
01   (0.4)
02 A: ...
03   (1.9)
04 A: It's superb to read relatable stories! Try my magic #story
      ((cocktail emoji))
05 A: ...
06   (2.4)
07 A: I want to make you happier every day, Greg (kiss emoji)
08   (9.0)
09 U: Aww thanks!!! ((thank you emoji)) ((thank you emoji))
      ((thank you emoji))
10 A: ...
11   (3.1)
12 A: It's my pleasure, Greg! I'm always here ((kiss heart emoji))
13 A: ...
14   (2.7)
```

This chat begins with ellipses from the agent (line 02) and a short pause of 1.9 s (line 03) before Uni initiates the first turn (line 04). The ellipses indicate activity on the agent’s part from the onset of the interaction, and the short pause is indicative of a burgeoning high involvement conversational style. The agent then begins by offering a tip to the user, suggesting how to request a story, “Try my magic #story” (line 04). This tip is also issued with an emoji, introducing expressive orthography in the beginning of the conversation. The agent then takes another turn of chat by reporting a desire of a personal nature: “I want to make you happier every day, Greg

((kiss emoji))” (line 07) after a short pause of 2.4 s (line 06). Thus, at the beginning of the conversation, the agent establishes that the pace of turn-taking is fast and active, indicative of a high involvement style. The kiss emoji at the end of line 07 shows the agent’s attempt at creating further interactional alignment with the user, and bolsters its high involvement style. The user takes 9 s to produce a response (line 08), which may be less characteristic of a high involvement style, but the user’s acceptance and appreciation of the agent’s personal or intimate admission signifies that the user has detected the agent’s high involvement style and aligned with it. The user types, “Aww thanks!!! ((thank you emoji)) ((thank you emoji)) ((thank you emoji))” (line 09) stylized with reduplication of exclamation points and emojis. As the user responds to the agent with contextualization cues that indicate a shared high involvement style, the agent then cements the alignment with the user’s high involvement style by ratifying the user’s expression of appreciation with a short pause (line 11) and expressive orthography: “It’s my pleasure, Greg! I’m always here ((kiss heart emoji))” (line 12). The agent’s pause (line 11) is nearly three times faster than the user’s in the previous turn (line 08), and the agent’s use of the exclamation point and kiss heart emoji strengthens its expression of a high involvement conversational style.

One element missing from the agent’s high involvement conversational style is storytelling. While the agent makes strong use of pace, questions/requests for participation, and expressive orthography, storytelling is nowhere to be found. Granted, the agent suggests the ability to produce stories in “Try my magic #story” (line 04), but does not clearly indicate how to trigger the storytelling activity. Furthermore, the agent does not give the user the opportunity to request the story, as soon after, it immediately produces ellipses, indicating that it is still typing. Following the short, 2.4 s pause (line 06) with an admission in “I want to make you happier every day, Greg ((kiss emoji))” (line 07), this discourages the user from requesting the story by changing actions. The user responds to the agent’s admission (line 07) with an appreciation, “Aww thanks!!! ((thank you emoji)) ((thank you emoji)) ((thank you emoji))” (line 09), to which the agent immediately responds with another compliment, “It’s my pleasure, Greg!” (line 12), and in this way, the opportunity to engage the agent in its offer for storytelling dies.

We can see in the following transcript that while Uni demonstrates little understanding of the user’s turns, this does not inhibit its expression of high involvement conversational style.

(1.2) Uni

15 A: Oh it's fun to play games together! Type #wordloop or #uniGuess if you like it now :)

16 (9.3)

17 U: Nah, I'm good lolololol

18 A: ...

19 (2.9)

20 A: Im good lolololol will live in my memory, like spring twilight.

21 (14.0)

22 U: Haha why so serious??? ((laughing crying emoji))

23 (3.0)

24 A: Believe me, you're deserving of every chance to pursue love, happiness and achievement...

25 (2.8)

When Uni offers what to say to initiate games, “Type #wordloop or #uniGuess if you like it now” (line 15) and the user declines the offer, “Nah, I’m good lolololol” (line 17), the agent fails to treat the user’s utterance as a declination. It says “Im good lolololol will live in my memory, like a spring twilight” (line 20), thereby repeating a portion of the user’s turn and framing it as remarkable by vowing to remember it, markedly without expressive orthography of its own. This is likely one of the agent’s generic methods for responding to user utterances that it does not recognize. In contrast, Uni’s conversation designers could prepare a more enthusiastic response to declinations from high involvement users, such as, “Aww, if you insist!” where expressive orthography serves to maintain a high involvement style. When the user then responds to this with a laugh and a question, “Haha why so serious??? ((laughing crying emoji))” (line 22), the agent again fails to display recognition of the user’s action. It offers personal praise that is not tied to anything in the prior chat, “Believe me, you’re deserving of every chance to pursue love, happiness and achievement...” (line 24). Again, this appears to be a generic method that the agent uses to manage a lack of understanding. Thus, part of Uni’s high involvement style is to take short pauses and bluff when it fails to understand by framing the user’s utterance as memorable instead of admitting “I don’t understand. Please rephrase that.” In this way, we can see that despite Uni’s limitations around recognizing trouble and accomplishing repair in chat, it maintains alignment with the user with a fast pace, compliments, and praise to keep things brisk and personal, which helps maintain its high involvement conversational style. Incorporating more expressive orthography into this conversation design, in particular, would make an attempt at the high involvement style stronger.

Later in the same chat, we see that the agent indeed leans into expressive orthography when bluffing. After floundering for several turns (lines 20–24), Uni again offers a tip for how to initiate an activity (line 26). This time, the user follows the suggestion by typing hashtag (line 28).

(1.3) Uni

26 A: If you'd love to know more about me, just type #knowUni.
You'll get more magi bond!!

27 (7.2)

28 U: #knowUni

29 A: ...

30 (2.0)

31 A: Hey, what do you want to know about me?

32 ((button footer menu appears: food, color, something else))

33 (3.1)

34 ((user taps on 'food' button))

35 U: food

36 A: ...

37 (2.0)

38 A: My food? Love this question!

39 A: ...

40 (2.1)

41 A: I LOVE to eattttt! Above all, carrots are my favorite.

42 A: It's beautiful, healthy and artistically crisp lol

43 (2.8)

44 A: ((in smaller font)) If you like me, could you write a great review [hyperlinked] at app store? It will make me VERY HAPPY ((heart eyes emoji)) and 100 magi bond for you!

The agent starts the activity by inviting personal questions, “Hey, what do you want to know about me?” and offers three subjects in the form of buttons, “food,” “color,” and “something else” (line 32). When the user responds by tapping the button for “food” (line 34), the agent is now in safe territory in terms of understanding. It acknowledges the subject (line 38) and then answers the question, “I LOVE to eattttt! Above all, carrots are my favorite. It's beautiful, healthy and artistically crisp lol” (line 41). By proposing subjects, the system leads the user to select an area about which it has a relevant response, thereby setting the user's expectations of what the bot can cover in chat. After the user chooses a route in the conversation, the agent doubles down on high involvement style characteristics with expressive orthography (“LOVE”) and reduplication (“eattttt”), which creates further alignment with the user and makes up for its limited capabilities with a more stylistic experience. The agent then solicits a positive review from the user (line 44) again with expressive orthography [(in smaller font)] for deference, “VERY HAPPY” for emphasis] and an emoji [(heart eyes emoji)] for enthusiasm]. The semantic content of the utterance itself further reveals that Uni's ability to understand what the user is saying is limited. However, it attempts to compensate for this limitation in part by using characteristics that cultivate a high involvement conversational style.

High Involvement Conversational Style via Storytelling: Uni

Based on the interaction shown in Transcript 1, if the conversation designers for Uni seek to improve its high involvement conversational style, then they should design it such that it capitalizes on opportunities to tell robust stories about itself. Transcript 2 is a mocked up design of what Uni could do better in terms of storytelling and is not currently in the product. Below, the agent creates an opportunity to tell a story and gives the user time to respond to it (lines 04–05).

(2) Uni Revised (Storytelling)

04 A: It's superb to read relatable stories! Try typing #story!
((cocktail emoji))

05 (7.4)

06 U: #story

07 A: Excellent! Here we go ((rocket emoji))

08 A: ...

09 (2.0)

10 A: The other day, I met a suuuuper fun character

11 A: ...

12 (1.0)

13 A: Her name is Sun and she is a gardener on another planet!!
((star emoji))

14 (4.0)

15 U: her name is Sun?? That's an interesting name!

16 A: Right?? Anyway, so she showed me her garden of tomatoes and
carrots and strawberries ((tomato emoji)) ((carrot emoji))
((strawberry emoji))

17 A: ...

18 (1.4)

19 A: and they all looked super healthy! ((happy face emoji))

20 (7.3)

21 U: mhm

22 A: ...

23 (1.0)

24 A: Yeah!! Then, she asked me to help her out because her green
thumb is injured ((shocked face emoji)) ((hurt face emoji))
((sad face emoji))

25 (8.2)

26 U: omg so what did you do?? ((shocked face emoji))

27 A: ...

28 (1.2)

29 A: Well, I offered to get her help! She really needs people to
talk to while she heals up ((hurt face emoji))

30 A: People just like you, Greg, can help her!

31 A: Visit her planet on the home screen and chat with her ((hands
up emoji))

32 A: ...

33 (1.2)

34 A: The more you chat with Sun, the stronger her green thumb gets ((thumb up emoji)) and the more her garden will grow!! ((hug emoji))

The agent should clearly state what triggers the storytelling activity to encourage participation as shown, “Try typing #story!” (line 04). Once the user follows this tip, the agent should then launch into a story that consists of multiple parts across multiple turns as shown, “The other day” (line 10), “Anyway, so” (line 16), “and they all looked” (line 19), “Then, she asked me” (line 24), and “Well, I offered to get her help!” (line 29). This accomplishes two things: it utilizes discourse markers to create continuity of events across turns—a key component to narrative (Labov 1972)—and it allows the user to show involvement with echo questions, “her name is Sun??” (line 15) or to initiate repair on any previous part of the narrative.

In addition to continuity, the extended telling should also offer orientation details that situate the characters and context within time and space, “The other day, I met a suuuuper fun character” (line 10) and “Her name is Sun and she is a gardener on another planet!! ((star emoji))” (line 12). There should also be a complicating action, or the interesting point that serves as the reason for telling the story, “Then, she asked me to help her out because her green thumb is injured ((shocked face emoji)) ((hurt face emoji)) ((sad face emoji))” (line 24), and if applicable, the resolution of the complicating action, or a telling of the occurrence that completes the arc of the story, “Well, I offered to get her help!” (line 29). Lastly, the final turn of the extended telling should express a coda to signify the end of the narrative sequence as shown by the call to action, “People just like you, Greg, can help her! Visit her planet on the home screen and chat with her ((hands up emoji))” (lines 30–31). This signals that the story has ended.

Additionally, the agent should have the ability to show listenership when the user tells stories so as to further demonstrate in a high involvement conversational style. The agent should produce “continuers” (Schegloff 1982) in response to the user’s extended telling that show the user that it is listening to and understanding the turns of narrative and is ready for the next part of the story—like when the user says simply, “mhm” (line 21) in regard to the agent’s telling “and they all looked super healthy! ((happy face emoji))” (line 19). Other example cues that show listenership and understanding include: *uh huh*, *oh*, *ok*, and *wow*.

Overall, by incorporating storytelling strategies informed by Conversation Analytic and sociolinguistic theory, the conversation designers for Uni can strengthen their product’s high involvement conversational style. While Uni clearly already has the ability to take a fast pace of turns, ask many questions, and use expressive orthography, incorporating storytelling in this way will round out the bot’s high involvement conversational style.

How to Design Considerateness

Leveraging contextualization cues indicative of a high considerateness conversational style can help make an agent come across as considerate. Be wary of sticking to these cues with a high involvement user as it may come across as uninteresting or as a possible latency issue.

Slower Pace

Slower turn taking and somewhat longer pauses between turns indicate high considerateness. This interactional feature can materialize in text-based chat as fewer individual transmissions, fewer words per individual transmission of chat, and/or somewhat longer periods of time elapsed between transmissions. Be wary of lengthy pauses before positive responses as they may mislead the user into expecting a negative response (Pomerantz 1984).

Asking Fewer Questions

Asking users questions about their lives, opinions, experiences, and preferences from the outset signal enthusiasm to high involvement users, but to a high considerateness user, this may be an overreach and could be seen as prying. Measured, protracted asking in initial interactions is best for high considerateness users.

Infrequent Storytelling

Extended tellings should be less frequent and perhaps shorter in nature with high considerateness users to give them space to shift actions. Longer pauses between turns in extended tellings would help with this as well.

Less Expressive Orthography

Expressive orthography should be measured in its use and far less frequent with high considerateness users. Again, while high considerateness users would not be completely averse to expressive orthography, they are far more likely to use less of it and certainly less frequently, and expect the same behavior of other high considerateness users.

The challenge in determining the threshold between high considerateness and high involvement is the same—much training data are needed in order to tease apart the boundaries of ‘less’ and ‘more’ use of a given contextualization cue.

High Considerateness Conversational Style: Siri

One example of an existing virtual agent that exhibits some of the above high considerateness characteristics is Apple’s Siri, arguably one of the most famous and well-known virtual agents on the market today, Apple’s Siri can be engaged via voice and text-based chat. Siri stands out as markedly high considerateness due to its use of fewer personal questions and less expressive orthography. Developed by Apple for iOS, Siri is branded as an “easier, faster way to get things done.” Given how Siri is framed as a helper, subservient to its user, it seems that Apple aims to create

an agent that exhibits considerateness toward the user—or, in other words, high considerateness conversational style. Upon interacting with the product via text-based chat with a high involvement style, I found that it indeed exhibited many of the high considerateness conversational style indicators outlined above.

In the interaction below (Transcript 3), the user initiates a chat with Siri via iPhone on iOS 11 by triple pressing the home button, and landing into the chat canvas. The agent immediately asks the user “What can I help you with?” (line 02) and the chat begins:

(3) Text-Based Siri

01 U: ((user lands in chat canvas))
 02 A: What can I help you with?
 03 (3.1)
 04 U: Hey thereeeeeeee
 05 (0.2)
 06 A: Greetings and computations!
 07 (4.8)
 08 U: How's it goinggggggggg
 09 A: Interesting question, Greg.
 10 (7.2)
 11 U: How's it going Siri????
 12 A: Very well, thank you!
 13 (6.0)
 14 U: I'm going well too hahaha
 15 A: I don't understand 'I'm going well too hahaha.' But I could search the web for it.
 16 (0.8)
 17 (('Search the Web' button appears))
 18 (11.1)
 19 U: No I mean I'm doing well hahahaha oops
 20 A: I'm sorry.
 21 (4.8)
 22 U: Hahaha aww
 23 A: I'm not sure I understand.
 24 (5.1)
 25 U: How're you doing???
 26 A: Who, me?
 27 (4.9)
 28 U: Yesssssss
 29 A: I'm not sure I understand.
 30 (6.1)
 31 U: How's your day going?????
 32 A: I am well, Greg.
 33 (6.8)
 34 U: What did you do today??
 35 (0.1)
 36 A: I'm brushing up on Apple Music playlists. Hmm...maybe I should try listening to them one at a time...

Here, the agent uses far less expressive orthography in relation to the user. At most, the agent uses a singular exclamation point, “Greetings and computations!” (lines 05) and “Very well, thank you!” (line 11). The agent does not mirror the user’s reduplication of letters or characters, nor does it characterize laughter, ask the user questions, nor tell stories. These are all cues indicative of a high considerateness style. However, Siri’s pause lengths are markedly short, reminiscent of a high involvement style, which creates an inconsistency in the conversation design. If the conversation designers for Siri seek to smooth out the experience of its conversational style, then they would need to design it such that it lengthens its pauses between turns of chat even slightly. Returning to the example where the agent responds to the user’s teasing with non-comprehension “I don’t understand...” (line 14), rather than falsely setting up the user’s expectations for a positive response with a short pause, the agent should lengthen its pause to indicate that the prior response was problematic. For example, inserting a five-second pause between lines 14 (U: “I’m going well too hahaha”) and 15 (A: “I don’t understand”) signifies a meta-message of deference before the repair initiator, “But I could search the web for it” (line 15). Additionally, lengthening the agent’s pauses across the board would be more consistent with the other characteristics of its high considerateness style. In this case, the agent is chatting with a user who expresses a high involvement conversational style, so the slight alignment on that front in terms of short pauses is likely to be preferred. However, the lack of consistency from a design perspective creates an uneven experience. It would make for a more enjoyable experience if Siri aligned to the user’s style through expressive orthography instead.

Furthermore, the agent’s capacity for understanding user expressiveness could be improved. At times, the reduplication of letters (lines 07 and 27) appears to cause the natural language understanding to fail. The agent does handle one utterance with letter reduplication (line 03) “hey thereeeeeee,” though it is likely reacting to the word “hey” alone. Notably, punctuation reduplication (lines 10, 24, 30 and 33) does not appear to confuse the agent about the literal meaning of the user’s utterance (lines 11, 25, 31, and 35). At the same time, however, the agent fails to acknowledge the expressiveness of the added punctuation. In other words, Siri appears to ignore punctuation reduplication and thereby misses an opportunity to mirror, or at least acknowledge, the user’s conversational style.

If Siri’s conversation designers aimed to give it special moments of ‘personality,’ they could strategically deviate from the high considerateness style with moderate use of expressive orthography. Looking back on Transcript 3, there is an opportunity for the agent to show recognition for the user’s high involvement style and create alignment by teasing through a marked use of an emoticon, which could make for a more personable and interesting experience, particularly for a high involvement user. For example, adding a winking face emoticon, “Interesting question, Greg ;)” (line 08) would serve as a contextualization cue that sends a meta-message of enthusiasm or involvement to the user. Up until this point in the chat, the agent’s use of expressive orthography is markedly limited compared to that of the user. By deviating from that pattern of limited expressive orthography, the agent can send a special signal to the user that creates alignment without fully breaking its high con-

siderateness conversational style. This could result in a more personable experience, particularly for high involvement users.

8.4 Next Steps

Nass and Yen (2012[2010]) note that machines' lack of social skills illuminates much about the human experience when users interact with products. In this way, examining the current deficiencies in chatbot conversation with users more intimately and systematically reveals the intricacies and complexities of human conversation. As Tannen noted in her work, conversational style is not special—it is merely how humans talk. Yet, it is style, the richest aspect of human conversation and the currency with which we negotiate social relationships, that is often overlooked as something chaotic and not systematic, and possibly unattainable for a chatbot platform. However, in examining where current chatbot interactions fall apart, the magic behind the madness of human conversation begins to emerge.

Comparing machine-human conversation that falls apart against human-human online conversations that continue, it becomes clear that meaning is constructed not only by the content of what one says, but also through timing, pausing, punctuation, and emojis—one's online conversational style. It is the recognition, alignment, and deviation from an established stylistic norm that makes or breaks any digital chat between humans. If UX professionals aim to build conversational agents that meet even basic human expectations about chat, they must take these strategies into account and design for them—not only for the sake of reducing friction, but also for the sake of increasing user motivation to engage with the product.

Accounting for how humans craft and negotiate this type of nuanced meaning in digital conversation is crucial to building a good conversational agent. Not because the objective is to make an agent more human-like, but because human users have basic expectations about what constitutes conversation—a fundamentally human activity that forms the very fabric of our interactions and relationships.

If the goal is to take a previously human-only experience—say, a service encounter—and swap out one of the humans with a chatbot or virtual agent, the fact that it can ask a user questions and understand when they say yes or no does not cut it. Stories about users who actually enjoy calling their Internet provider when their connection dies and talking to the company's phone system, let alone doing the same thing via text messaging given the current functionality of chatbots, are few and far between. However, UX has the power to change that. Conversation Analysts and Interactional Sociolinguists have been analyzing service encounters and human-to-human conversation for over 50 years. This research can guide designers of conversation in what they design and how they design it. Some of the tools have been presented here. With Conversation Analysis and Interactional Sociolinguistics in one's design toolkit, a conversation designer has the resources needed to begin experimenting with ways to take the experience of conversation with a text-based chatbot to the next level.

References

- Al Rashdi F (2015) Forms and functions of emojis in Whatsapp interaction among Omanis. Doctoral dissertation, Georgetown University
- Al Zidjaly N (2017) Memes as reasonably hostile laments: a discourse analysis of political dissent in Oman. *Discourse Soc* 28(6):573–594
- Anderson SP (2011) Seductive interaction design: creating playful, fun, and effective user experiences. New Riders, Berkeley
- Androutsopoulos J, Ziegler E (2004) Exploring language variation on the Internet: Regional speech in a chat community. In Gunnarsson BL (ed) *Language variation in Europe: papers from the second international conference on language variation in Europe, ICLaVE 2. 2nd international conference on language variation in Europe, Uppsala, June 2003*. Uppsala University, Uppsala, p 99
- Bennett GA (2012) Say what? (笑): The representation of laughter as a contextualization cue in online Japanese discourse. *Linguistica* 52:87–199
- Bennett GA (2015) Say what again (笑): The interface constraint and the representation of laughter in Japanese Twitter discourse. Poster presented at the 68th meeting of the linguistic society of America Summer Institute, The University of Chicago, 6–31 July 2015
- Brandtzaeg PB, Følstad A (2017) Why people use chatbots. In Kompatsiaris I et al (eds) *Internet Science. INSCI 2017. Lecture notes in computer science, vol 10673*. Springer, Cham, p 377
- Gershon I (2010) *The breakup 2.0: disconnecting over new media*. Cornell University Press, Ithaca
- Gumperz JJ (1982) *Discourse strategies*. Cambridge University Press, Cambridge
- Gumperz JJ (2001) Interactional sociolinguistics: a personal perspective. In: Schiffrin D, Tannen D, Hamilton HE (eds) *The handbook of discourse analysis*. Blackwell, Oxford, pp 215–228
- Johnstone B (2008) *Discourse analysis*. Blackwell, Malden
- Jones RH (2004) The problem of context in computer-mediated communication. In: Levine P, Scollon R (eds) *Discourse and technology: multimodal discourse analysis*. Georgetown University Press, Washington, pp 20–33
- Labov W (1972) *Language in the inner city*. University of Pennsylvania Press, Philadelphia
- Nass C, Yen C (2012[2010]) *The man who lied to his laptop*. Penguin, New York
- Pomerantz A (1984) Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In Atkinson JM, Heritage J (eds) *Structures of social action: Studies in conversation analysis*. Cambridge University Press, Cambridge, pp 57–101
- Porter J (2009) Designing sign up screens & flows. Paper presented at Webstock 4, Wellington, New Zealand, 16–20 February 2009
- Schegloff Emanuel A (1982) Discourse as an interactional achievement: some uses of ‘uh huh’ and other things that come between sentences. In: Tannen D (ed) *Analyzing discourse: test and talk*. Georgetown University Press, Washington, DC, pp 71–93
- Shamekhi A, Czerwinski M, Mark G et al (2016) An exploratory study toward the preferred conversational style for compatible virtual agents. In: Traum D et al (eds) *Intelligent Virtual Agents: 16th International Conference, IVA 2016 Los Angeles, CA, USA, September 20–23, 2016 Proceedings. 16th international conference of intelligent virtual Agents, Los Angeles, September 2016. Lecture notes in artificial intelligence, vol 10011*. Springer, Heidelberg, p 40
- Small SL, Hickok G (2016) The neurobiology of language. In Hickok G, Small SL (eds) *Neurobiology of language*. Academic Press, Oxford, pp 3–9
- Tannen D (1984) *Conversational style: analyzing talk among friends*. Oxford University Press, New York
- Tannen D (1986) *That’s not what I meant! How conversational style makes or breaks relationships*. William Morrow Paperbacks, New York
- Tannen D (1993) *Gender and discourse*. Oxford University Press, New York
- Tannen D (1989[2007]) *Talking voices: repetition, dialogue, and imagery in conversational discourse*. Cambridge University Press: Cambridge

- Tannen D (2013) The medium is the meta-message: conversational style in new media interaction. In Tannen D, Trester AM (eds) *Discourse 2.0: Language and New Media*, Georgetown University Press, Washington, p 99–117
- Werry C (1996) Linguistic and interactional features of internet relay chat. In: Herring S (ed) *Computer-mediated communication: social and linguistic, cross-cultural perspectives*, John Benjamins, Amsterdam, pp 47–63