

# Modeling the Rhetoric of Human-Computer Interaction

Iris Howley, Carolyn Penstein Rosé,

Carnegie Mellon University,  
5000 Forbes Ave. Pittsburgh, PA USA  
{ihowley, cprose}@cs.cmu.edu

**Abstract.** The emergence of potential new human-computer interaction styles enabled through technological advancements in artificial intelligence, machine learning, and computational linguistics makes it increasingly more important to formalize and evaluate these innovative approaches. In this position paper, we propose a multi-dimensional conversation analysis framework as a way to expose and quantify the structure of a variety of new forms of human-computer interaction. We argue that by leveraging sociolinguistic constructs referred to as *authoritativeness* and *heteroglossia*, we can expose aspects of novel interaction paradigms that must be evaluated in light of usability heuristics so that we can approach the future of human-computer interaction in a way that preserves the usability standards that have shaped the state-of-the-art that is tried and true.

**Keywords:** computational linguistics, dialogue analysis, usability heuristics

## 1 Introduction

As computing continues to grow more ubiquitous and new and innovative types of interaction possibilities emerge in a variety of settings far away from the desktop, the question of how to formalize and evaluate these new forms of interactions becomes an increasingly important question. The relationship between man and machine will evolve in unpredictable ways, and as computers become more situated in daily life we will need to be able to learn from the new roles both entities adopt when communicating with one another. Frameworks formalizing the relationship between user and novel interface can be leveraged towards standardizing, testing, and evaluating with respect to usability heuristics. In this position paper, we begin by proposing a multi-dimensional framework for analysis of human-human conversations for the purpose of better understanding these social relationships between humans and computers. We then discuss how this might eventually generalize to a wider variety of interaction styles between humans and computers.

## 2 Motivation

As technology develops, computers are becoming more and more autonomous, resulting in mixed-initiative interactions, a situation in which entities (both human and computer) in the interaction contribute what is most appropriate at the most appropriate time [1]. Additionally, more systems developers are adopting the approach of adjustable autonomy, where the system can automatically modify its level of independence and control in order to help the user without becoming overly intrusive [2]. Determining when the agent should pass control to the user is already an established problem within the area of adjustable autonomy research, and as such, various methods have already been established to determine the successes and failures of these systems [3]. For example, Scerri and colleagues [4] introduce a generalized transfer-of-control strategy which includes two machine learning approaches, namely C4.5 and Markov decision processes. These approaches produce control structures that are rule-based (i.e. “if the department head is not at the meeting and it is a Monday, keep control”) or constraint-based (i.e. expected quality of the agent’s decision) [4]. However, these approaches are limiting in that they can be thought of as treating interaction overly mechanistically and thus ignoring the social aspects of the interaction. Research has already established that humans behave in social ways with systems that are not intentionally social, such as in Forlizzi’s work [5] where humans establish social relationships with their non-social, domestic vacuuming robots. Once humans begin to engage socially with computational systems, they begin to adopt a different set of expectations and to orient towards these expectations in their behavior. These behaviors and expectations then lead to differences in the balance of control and responsibility than what is typical in current more typical interaction paradigms. Thus, computational systems that potentially elicit this type of engagement with humans need to not just act accurately, but act with proper consideration to their social role within the interaction.

The majority of work on these approaches happens within the sphere of human-robotic interaction, but as personal computers become more embedded and intelligent, mixed-initiative and adjustable autonomy systems will become a reality on a daily basis. This emerging reality leaves us with important questions to answer. When the computer performs actions on its own initiative, without being commanded by the user, how will we evaluate what effect that action has on the system’s ability to reduce user memory load, provide feedback, or prevent errors [6]? From a designer’s perspective, how will we formalize what happens in these new interactions so that we may evaluate how the properties of the new interactions affect usability of the system?

*Jarvis:* Test complete. Preparing to power down and begin diagnostics.

*Stark:* Yeah. Tell you what. Do a weather and ATC check.

*Stark:* Start listening in on ground control.

*Jarvis:* Sir, there are still terabytes of calculations needed before an actual flight is...

*Stark:* Jarvis! Sometimes you got to run before you can walk.

- “Iron Man” Film Script, 59:10.

As an example, the dialogue segment above shows a sample piece of fictional interaction between a somewhat autonomous computer (Jarvis) and its user (Stark). Traditional usability standards can still apply in such situations. For example, the first line where the user is being kept current with the status of the machine is exemplary of the system providing good feedback. Later, when Jarvis attempts to warn Stark before he requests something exceedingly taxing on the system could be considered approaching the prevention of an error with serious consequences (i.e., system destruction). However, we see that the user ignores the computer's warning, and the computer lets him do so. Typically offering users the option of a manual override is considered a positive usability feature. In the case of a conversational interaction where negotiation over control and responsibility is more subtle than a manual override button, it's not always apparent where the transfer of control and responsibility should and does take place. Thus, despite following some basics of usability design, there is still a serious breakdown that occurs shortly after this event in the script. This suggests a possible tension between the usability heuristics as computer systems take on more intelligent roles and interactions with them become more of a dance. In order to determine what about this interaction leads to these positive and negative assessments of usability heuristics we formalize the stylistic aspects of the interaction using a conversational analysis framework.

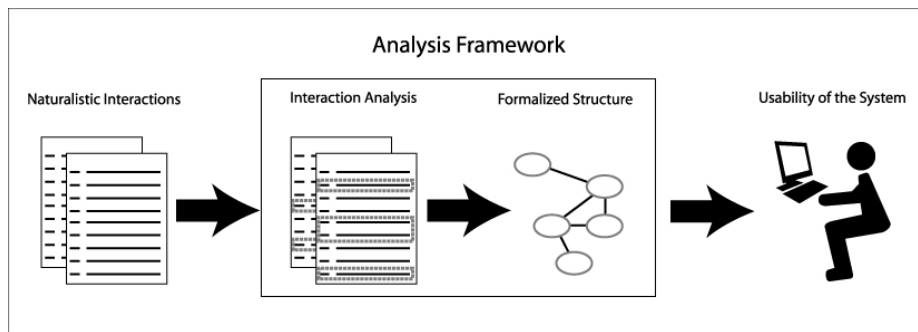
Past work [7] has already examined how lessons from human-human interaction in the social sciences can help inform the design of ubiquitous computing systems. In a similar spirit, in this paper we propose using methods from human-human interaction in sociolinguistics to develop a methodology to help evaluate emerging interaction paradigms. We can apply past work on human-human dialogue analysis to formalize features of human-computer interactions, perhaps even where they are not explicitly conducted in natural language, in order to investigate the usability of these new systems. The process we envision for formalizing the structure of interactions and applying to evaluating the usability of systems is shown in Figure 1.

For the past four years our research group has worked on a computational framework for the analysis of leadership in group conversations [8], [9], [10], [11]. The work we have done has mainly been applied in the area of Computer-Supported Collaborative Learning. Evidence from our work in this area has shown that leadership is an important multi-faceted construct for investigating social positioning within a group learning context. We examine the role of leadership in conversation and social positioning through three dimensions: displays of reasoning (i.e. transactivity), contribution authoritativeness (i.e., Negotiation), and contracting/expanding for additional viewpoints (i.e., Heteroglossia). While previously this framework was applied to human-human communication, as humans were the most embedded participants in a conversation, increasingly, computers are more involved as active participants in interactions.

The majority of interactions in traditional computing environments are not conducted conversationally. However, we argue that our framework provides a conceptualization that can be adapted for other purposes. For example, the Negotiation framework embodies the idea that a display of needing validation or requiring information is a sign of non-authoritativeness. On a simple level, detection of uncertainty in a user's interaction with the system could be seen as operating on this dimension, for example, timing that indicates hesitation, or perhaps a style of

touch that indicates tentativeness. The Heteroglossia framework might relate more to the quantity of constraints indicated by users in the preferences. Users who make extensive use of customization options could be seen as behaving more authoritatively.

Users who are authoritative in a functional way will develop creative ways of using the potential offered by the environment. This shows an awareness of what the system offers while not abdicating control to the system. This multi-dimensional framework would suggest that in a mixed-initiative environment, we want to see users achieving their own goals while taking advantage of what the system has to offer. The triangulation enabled through an integration of the three dimensions of this framework allows for identification of users who are adopting a dysfunctional authoritative stance, and ideally, support for adopting a more functional and effective stance within the environment.



**Figure 1.** Shows the process of how our analysis scheme allows us to formalize aspects of the structure of more naturalistic man-machine interactions so that we can systematically investigate how the properties of such interactions affect the usability of the system.

## 4 The SOUFFLÉ Framework

In this paper, we focus on two dimensions of the SOUFFLÉ Framework. We consider the idea of leadership from two directions: first in terms of how authoritative a speaker presents herself, and second, in terms of how receptive a speaker is to the leadership of others. In terms of how authoritatively a speaker presents herself, we adapt two constructs from the field of systemic functional linguistics, namely Martin and Rose's Negotiation Framework [12] and Martin and White's operationalization of heteroglossia [13]. In the Negotiation Framework, authoritativeness is demonstrated by making a contribution to a discourse that is not offered as an invitation for validation from another group member. In contrast, within the Heteroglossia Framework, assertions framed in such as way as to acknowledge that others may or may not agree, are identified as heteroglossic. In representing these two dimensions separately we are able to identify those rare leaders who can present their ideas as standing on their own without denying others their own voice. Each of these

dimensions presents a different take on how authoritatively a speaker is positioning himself within an interaction, which provides a rich view of leadership that allows us to see that a speaker's role adoption may not be complete, or may be dysfunctional in some way. Further details on the dimensions of our framework can be found in [11].

#### **4.1 The Negotiation Framework**

The Negotiation Framework is a measure of authoritativeness where authority is demonstrated by making a contribution to a conversation that is not meant to be validated by another group member. Past work has shown a relationship between authoritativeness and group self-efficacy within computer-supported collaborative learning tasks, speaking to the importance of authoritativeness in group functioning [14]. As part of the process of exploring this authoritativeness relationship between speakers, we have developed a coding manual so we may better identify the ebb and flow of social power within a conversation. The Negotiation Framework includes four core moves, and two secondary moves:

- K1 (Primary Knower), in which the speaker considers herself to be the primary authority on a given (expressed) piece of knowledge
- K2 (Secondary Knower), when the contributor asks for knowledge from someone of higher authority
- A1 (Primary Actor), for contributions that display that one can perform a particular action.
- A2 (Secondary Actor), when instructing someone else to do an action, allowing the other person to either perform the action or reject the request.
- ch (Challenge), in which a speaker rejects the authority of the previous speaker to make the previous move
- o (Other), which encapsulates all other moves that do not fit in the five described above

For our purposes, "Primary Knower" and "Secondary Actor" moves are considered more authoritative (with respect to social relationships), while "Secondary Knower" and "Primary Actor" moves display less authoritativeness. As such, to compute a meaningful ratio for the authoritative moves, the formula would be:  $(K1 + A2) / (K1 + K2 + A1 + A2)$ . It should be noted that these moves are only applicable for situations where contributions are being directed at other members of the conversation. There are several more complications to the coding process using the Negotiation Framework, a discussion of which can be found at [11]. We will explore several examples in the next section which should clarify these concepts considerably.

#### **4.2 The Heteroglossia Framework**

The Heteroglossia framework is operationalized from Martin and White's theory of engagement [13], and here we describe it as identifying word choice that allows or restricts other possibilities and opinions. This creates a rather simple divide in possible coding terms for contributions:

- Heteroglossic-Expand (HE) phrases tend to make allowances for alternative views and opinions (such as “She *claimed* that usability heuristics are great.”)
- Heteroglossic-Contract (HC) phrases attempt to thwart other positions (such as “She *demonstrated* that usability heuristics are great.”)
- Monoglossic (M) phrases make no mention of other views and viewpoints (such as “Usability heuristics are great.”)

In some situations hedging words such as “probably” and “I think” also have a similar distancing effect as either of the Heteroglossic codes, despite the fact that the speaker may simply be unable to commit to a statement due to lack of knowledge or any other reason. Our research group has examined giving dialog agents heteroglossic, monoglossic, and neutral language in an idea generation task, and found that dialog agents with heteroglossic language result in the greatest idea generation productivity in a group task [15]. Because the level of heteroglossia an agent embodies has important effects on user productivity, we chose to focus on it in our sample analyses below.

## 5 Application to More Informal Interactions

User actions can be interpreted with respect to authoritativeness and heteroglossia and in this section we will go into the specifics of how looking at these actions and interactions between computers and humans can inform interaction design.

As part of our work, our research group has collected a large quantity of datasets detailing interactions between multiple humans, between a human and a computer, and between multiple humans and a computer. This collection of discussion datasets reflect what the current state of computer-human conversation interactions are currently, but what do we know about where these conversational interactions are headed? How can we use our knowledge of how the social relationships between humans and computers are changing to inform our design of interaction systems?

Here we introduce some snippets of script from the movie, “Iron Man”. We specifically selected conversational interactions between the lead character, Tony Stark (who is the Iron Man) and his home computer, Jarvis. We believe these interactions are informative as idealized interactions between man and machine, and intend to use an analysis of the script to inform future interactive systems without limiting ourselves solely to current dialogue systems. The authors would like to stress that we in no way intend to imply that the selected scenes between Jarvis and Stark from the “Iron Man” movies is indicative of future user interfaces. It is simply an appropriate example in which to show the features and flexibility of our framework for addressing the kinds of questions that arise when we consider the potentials of interactions enabled through emerging technologies in artificial intelligence and language technologies.

Table 1 is a brief interaction between Stark and his sentient computer, Jarvis, from near the beginning of the movie. We see in lines 1-4 that Stark is interacting with Jarvis in a rather authoritative way (i.e., commanding action in lines 1 and 3). These direct commands can be translated to non-dialog interaction with a system as clicks of a mouse. “Give me an exploded view” could be the equivalent of using menus to

change the view of a particular system. In the film, Jarvis performs the actions commanded to him, and so we see how these commands place Stark in a considerably more authoritative role than the computer. Stark's commands, coded within the Negotiation framework as A2, reflect a formalized version of this authority that can later be used to evaluate the interaction design.

**Table 1.** A short interaction between Stark and his computer from the film, "Iron Man" at time point 10:35.

Line	Text	Neg.	Heterog.
1	<b>Stark:</b> Give me an exploded view.	A2	M
2	<b>Jarvis:</b> The compression in cylinder three appears to be low.	K1	HE
3	<b>Stark:</b> Log that.	A2	M
4	<b>Stark:</b> I'm gonna try again, right now.	A1	M
5	<b>Stark:</b> Hey, Butterfingers, come here.	A2	M
6	<b>Stark:</b> What's all this stuff doing on top of my desk?	K2	M

The Heteroglossia analysis also reflects the commanding nature of Stark's requests. We see in line 2 that Jarvis allows for other possibilities with the word "appears", rather than other options such as "is reporting" instead. In comparison to Jarvis' heteroglossic statement, Stark's words are entirely monoglossic commands, not allowing other options. In lines 5 and 6 Stark's conversational style with the computer changes. He sarcastically calls the computer "Butterfingers" and appears irked at the machine as he asks the question in line 6. Stark still seems in a more authoritative stance than Jarvis with his A2 Negotiation command, but the interaction style has changed and the Negotiation codes reflect this. Now, instead of simply telling the system to perform an action, Stark asks the computer an open-ended question, denoted by the K2 Negotiation code. Line 6 shows us that the computer is potentially capable of helping its user in confusing situations.

Similarly, Table 2 shows a small portion of conversation between Stark and his conversational computer, Jarvis, from a much later point in the film. The tone has changed from the previous conversation segment yet again, with both contributors taking on mostly authoritative roles (K1 and A2 codes). In line 9, Stark has taken to adjusting his language to be less commanding. Instead of demanding "Open a new project file", Stark states that he'd "like to open a new project file". Similarly, in lines 11 and 12 we see in the Heteroglossia coding that Stark has opened up his statements to other possibilities. Words such as "actually" and particular phrasing like "why don't we" show that Stark is aware of other options. This is behavior seen throughout the script being performed mostly by Stark, with Jarvis rarely using heteroglossic statements.

**Table 2.** A clip of script from the movie "Iron Man" occurring around time 54:05.

Line	Text	Neg.	Heterog.
------	------	------	----------

7	<b>Stark:</b> Jarvis, you up?	K2	M
8	<b>Jarvis:</b> For you, sir, always.	K1	M
9	<b>Stark:</b> I'd like to open a new project file, index as Mark Two.	A2	M
10	<b>Jarvis:</b> Shall I store this on the Stark Industries Central Database?	A2	M
11	<b>Stark:</b> Actually, I don't know who to trust right now.	o	HC
12	<b>Stark:</b> Till further notice, why don't we just keep everything on my private server?	A2	HE
13	<b>Jarvis:</b> Working on a secret project, are we, sir?	K2	M
14	<b>Stark:</b> I don't want this winding up in the wrong hands.	K1	M
15	<b>Stark:</b> Maybe in mine, it can actually do some good.	K1	HE

Unlike the previous example where Stark's monoglossic commands can be translated into direct interface manipulations, heteroglossic statements cannot. In this situation, our framework can formalize this emerging interaction style and allow us to evaluate different approaches to handling it. We can still examine this interaction from a current usability heuristics standpoint. For example, Jarvis' line 10, "Shall I store this on the Stark Industries Central Database?" is likely an obvious attempt at reducing the user's cognitive load [6]. Suggesting a file-save location may save Stark from spontaneously recalling the most relevant save location from memory. Furthermore, Jarvis' awareness of the privacy concerns related to the document in discussion can suggest putting effort towards merging security and privacy with human-computer interaction, a field further discussed in [16].

Table 3 shows another conversation between Stark and Jarvis from a later point in the movie. Jarvis' lines in this segment are fairly typical of appropriate system feedback in an attempt to prevent errors, yet again (this time with a better ending than in the example from the introduction). The importance of the warning Jarvis communicates requires a greater level of persistence, until Stark acknowledges the warning in line 24 and essentially minimizes it. This exchange is reflected in the Negotiation coding with an A2-o repetition, with Stark once again taking an authoritative role. Stark's commanding position is reflected in the Heteroglossia coding as well, with entirely monoglossic codes. Since this particular interaction between Stark and his computer ended successfully, we could use these codes to examine what was unique in this interaction to make it successful. Was it the A2-o pattern, or the entirely clear-cut monoglossic statements that led to the success of this interaction? Are these desirable interaction styles?

**Table 3.** A piece of script from the movie "Iron Man" occurring later in the movie.

Line	Text	Neg.	Heterog.
16	<b>Stark:</b> Take me to maximum altitude.	A2	M
17	<b>Jarvis:</b> With only 15% power, the odds of reaching	ch	M



	that...		
18	<b>Stark:</b> I know the math! Do it!	A2	M
19	<b>Jarvis:</b> Thirteen percent power, sir.	o	M
20	<b>Stark:</b> Climb!	A2	M
21	<b>Jarvis:</b> Eleven percent.	o	M
22	<b>Stark:</b> Keep going!	A2	M
23	<b>Jarvis:</b> Seven percent power.	o	M
24	<b>Stark:</b> Just leave it on the screen!	A2	M
25	<b>Stark:</b> Stop telling me!	A2	M

---

We see in these episodes a difference in footing between the user and machine that is indicated through differences in the codes assigned by the conversational analysis framework. As we then find differences in desirability between these modes of communication and collaboration between man and machine, we can use this framework to make precise what aspects of the interaction account for these differences, and thus use this analysis approach as a step towards informing design.

## 6 Conclusion

Our analyses from the previous section shows that our framework can formalize several of the social dimensions occurring in a more informal interaction between a human and a computer. These formalizations can then be used for evaluation of systems or to inform future designs where the relationship between human and computer is less obvious. Basing our frameworks on well-developed sociolinguistic theories that we formalize provides us with a powerful lens for challenging these new interaction techniques in the light of issues such as authority and responsibility.

In conclusion, we have shown that these frameworks originally developed for human-human discussion can be extended to human-computer dialog, inform usability for more general human-computer interaction, and it can be argued that its utility reaches even further. We may be able to use the frameworks to explore usability not just for more informal interactions and dialog systems, but possibly for even more naturalistic methods of interacting with computers, such as via gestures. While we have not, as of yet, explored how our frameworks may inform usability within gesture interactions, it seems plausible that one could detect more and less authoritative movements or more and less expansive gestures. For example, quick movements and tense body language could imply authority. The flexibility of our framework to formalizing emerging interaction systems is one of the strengths of the SOUFFLÉ framework, which allows us to expose the structure in human-computer interaction for current systems and systems yet to be invented.

## References

1. Goodrich, M.A., Schultz, A.C.: Human-Robot Interaction: A Survey. In: Foundational Trends of Human-Computer Interaction, 1, 3 (January 2007), 203-275.
2. Scerri, P., Pynadath, D., Tambe, M.: Adjustable Autonomy in Real-World Multi-Agent Environments. In: Proceedings of the Fifth International Conference on Autonomous Agents (AGENTS '01). ACM, New York, NY, USA, pp. 300-307 (2001).
3. Fleming, M., Cohen, R.: A Utility-Based Theory of Initiative in Mixed Initiative Systems. In: The IJCAI-01 Workshop on Autonomy, Delegation, and Control: Interacting with Autonomous Agents (2001).
4. Scerri, P., Pynadath, D.V., Tambe, M.: Why the Elf Acted Autonomously: Towards a Theory of Adjustable Autonomy. In: Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems: part 2 (AAMAS '02). ACM, New York, NY, USA (2002).
5. Forlizzi, J. How robotic products become social products: An ethnographic study of cleaning in the home. In: Proceedings of HRI'07, 129-36, (2007).
6. Nielson, J.: Usability Engineering. Academic Press, San Francisco (1993).
7. Bellotti, V., Back, M., Edwards, W.K., Grinter, R.E., Henderson, A., Lopes, C.: Making Sense of Sensing Systems: Five Questions for Designers and Researchers. In: Proceedings of the Conference on Human Factors in Computing Systems (CHI 2002). ACM Press, Minneapolis, MN, USA, pp. 415-422, (2002).
8. Joshi, M., Rosé, C.P.: Using Transactivity in Conversation Summarization in Educational Dialog. In: Proceedings of the SLATE Workshop on Speech and Language Technology in Education (2007).
9. Rosé, C.P., Wang, Y.C., Cui, Y., Arguello, J., Stegmann, K., Weinberger, A., Fischer, F.: Analyzing Collaborative Learning Processes Automatically: Exploiting the Advances of Computational Linguistics in Computer-supported Collaborative Learning. Submitted to: The International Journal of Computer-Supported Collaborative Learning 3(3), pp. 237-271 (2008).
10. Ai, H. Sionti, M., Wang, Y.C., Rosé, C. P.: Finding Transactive Contributions in Whole Group Classroom Discussion. In: Proceedings of the International Conference of the Learning Sciences (2010).
11. Howley, I., Mayfield, E., Rosé, C. P.: Linguistic Analysis Methods for Studying Small Groups. In: Cindy Hmelo-silver, Angela O'Donnell, Carol Chan, & Clark Chin (eds.) International Handbook of Collaborative Learning, Taylor and Francis, Inc. (to appear).
12. Martin, J.R., Rose, D.: Working with Discourse: Meaning Beyond on the Clause. Continuum (2007).
13. Martin, J. R., White, P.R.R.: The Language of Evaluation: Appraisal in English. Palgrave: Macmillan (2005).
14. Howley, I., Mayfield, E., Rosé, C.P.: Missing Something? Authority in Collaborative Learning. To appear in: Proceedings of the 9th International Conference on Computer-Supported Collaborative Learning (2011).
15. Kumar, R., Beuth, J.L., Rosé, C.P.: Conversational Strategies that Support Idea Generation Productivity in Groups. To appear in: Proceedings of the 9th International Conference on Computer-Supported Collaborative Learning (2011).
16. Johnston, J., Eloff, J.H.P., Labuschagne, L.: Security and Human Computer Interfaces. In: Computers & Security, 22(8), pp. 675-684 (2003).