Time is On My Side: Dynamics of Talk-Time Sharing in Video-chat Conversations

KAIXIANG ZHANG, Cornell University, USA
JUSTINE ZHANG, University of Michigan, USA
CRISTIAN DANESCU-NICULESCU-MIZIL, Cornell University, USA

An intrinsic aspect of every conversation is the way talk-time is shared between multiple speakers. Conversations can be balanced, with each speaker claiming a similar amount of talk-time, or imbalanced when one talks disproportionately. Such overall distributions are the consequence of continuous negotiations between the speakers throughout the conversation: who should be talking at every point in time, and for how long?

In this work we introduce a computational framework for quantifying both the conversation-level distribution of talk-time between speakers, as well as the lower-level dynamics that lead to it. We derive a typology of talk-time sharing dynamics structured by several intuitive axes of variation. By applying this framework to a large dataset of video-chats between strangers, we confirm that, perhaps unsurprisingly, different conversation-level distributions of talk-time are perceived differently by speakers, with balanced conversations being preferred over imbalanced ones, especially by those who end up talking less. Then we reveal that—even when they lead to the same level of overall balance—different types of talk-time sharing dynamics are perceived differently by the participants, highlighting the relevance of our newly introduced typology. Finally, we discuss how our framework offers new tools to designers of computer-mediated communication platforms, for both human-human and human-AI communication.

CCS Concepts: • **Human-centered computing → Collaborative and social computing design and evaluation methods**.

Additional Key Words and Phrases: conversations; interaction patterns

1 Introduction

In a conversation, talk must be somehow distributed among the speakers. Some people talk more than others, perhaps remarkably so—thus the complaint that one is being talked over, or the worry that one has talked too much. In a classic study of talk at a dinner party, Tannen [88] examines many aspects of the guests' conduct; among them, the "high-involvement" behaviors of some guests— who talk a lot and leave little room for interjection—lead others to report feeling "dominated." Tannen's analysis additionally traces how such feelings of fitting in or getting left out are produced, as the dinner conversation moves from back-and-forth banter to focused attention on a single storyteller. In this way, the distribution—and constant renegotiation—of talk-time in a conversation can meaningfully shape the speakers' experiences.

In this paper, we explore how talk-time is distributed between speakers in video-chat conversations. We go beyond thinking of this distribution as a static, conversation-level characteristic—i.e., the extent to which some speakers talk more than others—and instead conceive of it as the result of a *dynamic* process that develops throughout the entire duration of the interaction. As we will show, attending to the dynamics is important for revealing potentially consequential distinctions. A perfectly balanced conversation can be comprised of rapid-fire back-and-forth exchanges or of alternating monologues; accordingly, a broad sense of interactional equity may be modulated by differing experiences of flow or engagement.

We propose a computational framework to examine the dynamics of talk-time sharing. Our approach builds off of a simple measurement: for a given span of conversation, we quantify the fraction of time taken by each speaker. Going span-by-span, we represent each conversation in terms of how the distribution of talk-time changes as the conversation progresses. Based on these representations, we then derive a space of possible talk-time sharing dynamics structured by several intuitive axes of variation.

To demonstrate its use, we apply our framework to a collection of over 1,500 video-chat conversations between pairs of strangers [63]. We examine how different types of talk-time sharing dynamics correspond to speakers' post-hoc assessments of how their conversations went. Perhaps unsurprisingly, we find that conversations that are overall more balanced in terms of the distribution of talk-time are preferred. However, we reveal striking differences in speakers' perceptions when comparing different dynamics that lead to such balanced conversations. In particular, balanced conversations where speakers engage in back-and-forth interactions are perceived as being significantly less enjoyable than those where speakers alternate in dominating the talk-time. Such distinctions, alongside others we observe, highlight the need for not only analyzing static snapshots of the conversation-level distribution of talk-time, but also the lower-level dynamics that lead to them.

We additionally use our framework to perform an exploratory analysis of the relation between talk-time sharing dynamics and speakers' characteristics. We find that speakers are consistent in their talk-time sharing behavior across multiple conversations, and that this consistency can in part be explained by age and gender demographics. Although exploratory and limited to a single domain, our analyses suggest potential ways for systematic inequities in the way speakers from different groups interact with each other to show up in conversations.

We also show how our framework can be used to surface meaningful conversation dynamics in other domains. While our analysis primarily focuses on the video-chat data, as an illustrative example, we apply the framework to a collection of Supreme Court Oral Arguments—a scenario that presents notable contrasts to the video-chat setting, since it involves multiple speakers with pre-determined roles interacting face-to-face. To encourage broader applications, we distribute the code that implements our framework, together with demonstrations on the video-chat and Supreme Court datasets.

From a practical standpoint, we argue that researchers should attend to the dynamics of talk-time sharing—beyond static qualities like the degree of conversation-level balance—when analyzing or designing for computer-mediated conversations. We offer our general framework as a tool for examining such dynamics and extending a design space of possible interventions that could potentially encourage more equitable [16, 43, 47, 48, 92] and enjoyable [30] interactions. In light of the increasing prominence of human-AI interactions, our framework also surfaces an additional dimension to consider when assessing the (in)capability of AI agents to have human-like conversations [15, 64, 86, 98].

From a theoretical standpoint, our paper considers a central preoccupation of existing scholarship on interaction: how conversation happens [15, 73, 79, 95, inter alia]. A wide range of past work in

fields like linguistics and sociology has focused on such mechanics as taking turns [49, 66, 85], telling stories [38, 51, 88], and bringing conversations to a close [72]. Our framework contributes a novel perspective on the dynamic distribution of talk, and a way of analyzing large-scale conversational corpora through that lens. In particular, while past studies have examined the interplay between regimes of talk where one speaker leads versus where both equally contribute [15, 19, 23], we offer a way of holistically modeling conversations in terms of such regimes so that they can be systematically compared to each other.

2 Background

Our work focuses on the mechanics of how conversation happens. In this way, we share similar interests as past work from fields like psychology, sociology, and linguistics that examines how people take turns [49, 66, 85], sequence actions [9, 70], address communicative troubles [71], engage in interactional modes like storytelling [2, 12, 15, 38], or start or end conversations [67, 72]—alongside numerous other aspects. Some work more explicitly engages with what people say—e.g., what feelings are expressed or what outcomes are arrived at (see Yeomans et al. [95] for a survey). Such factors are somewhat out of scope for our study, though we draw on them when interpreting our dynamics of interest. Rather, we engage with the presumption—shared by a diverse range of paradigms from conversation analysis [35, 73] to computational linguistics [15, 26, 27]—that it's worth accounting for the things people do that make conversation (dis)orderly or (in)coherent. Conversation analysis in particular attends to how coherence is *achieved* by speakers as an interaction progresses [35, 73]; guided by this sensibility, we make the *constant renegotiation* of talk-time distribution a central focus of our approach.

One recurring concern in such research is how talk is allocated. Numerous studies focus on turn-taking: who speaks next, and when do they start [49, 66, 69]? In investigating how speakers "orient to" the norm of "one speaker at a time," some work also considers violations of this norm, like interruptions [17, 25] or silences [33, 44, 65]. Other work goes beyond individual turns to examine who has the floor [18, 24], who's in control [39, 55, 93], or who's at the center of attention [26, 27]. This perspective abstracts away from ambiguities relating to what counts as a turn: an individual utterance could be construed as an interruption, a backchannel, or a meaningful change in who's speaking [24, 49].

Our work adds to these accounts of the distribution of talk. We consider broader spans of conversation, as opposed to individual utterances or turns, placing our conceptualization closer to those about control or floor. Our work is not the first to do this—other accounts have drawn a distinction between "chunks" (where one speaker takes the floor and dominates for an extended period) and "chats" (where the floor may be jointly developed), documenting their distributions and the ways they transition between one another [15, 18, 19, 23]. However, we go on to propose ways of comparing between different types of conversations, based on these characteristics. This allows us to go beyond corpus-level claims about the composition of talk—where such studies have generally focused—to provide a systematic account of the diversity of conversational dynamics within a corpus.

Our computational approach and our level of abstraction present certain compromises. Our framework should not be interpreted as an accurate model of ideas like control or floor—which concern speakers' subjective impressions [18, 24]—so much as a means of methodically surfacing interactional patterns that could point to these ideas. Our coarse focus on talk-time ignores the intricate and contextual interpretations of detail that are characteristic of approaches like conversation analysis [35, 68, 79]. Instead, following other writing on the complementary perspectives offered by quantitative and qualitative methods (e.g., [1, 60]), we see our approach as providing a birds-eye view over a space of conversational dynamics (see [96] for an analogue in a social-media

setting). By systematically applying a simple measure—share of talk-time—across subsequent spans of interaction, we are able to produce holistic representations of conversations, organize them in an intuitively structured space, and make claims about their relative prevalence in particular contexts. As we discuss in Section 6, there are many ways for future, finer-grained approaches to elaborate on the scaffolding we've set up.

Relation to speakers and their experiences. How do different ways of allocating talk relate to speakers' impressions of a conversation? Past studies often approach this question in terms of whether or not interactions are balanced, or whether certain people contribute larger shares of talk. Across multiple settings, researchers have drawn connections between balance and other factors like task performance [6, 56, 59] and group dynamics [20, 50, 52, 62]. Numerous studies have also examined the relationship between gender and the distribution of talk [4, 5, 37, 52], often drawing connections to the unequal levels of power conventionally held by men versus women (e.g., [5]).

Past work generally conceives of balance as a good thing, especially when speakers are on equal footing: per Guydish et al. [29] and Guydish and Tree [30], conversational balance could reflect that interlocutors have complementary, reciprocal behaviors, and could result in decreased social distance. Per Lakoff [46] and Tannen [88], dominating the conversation could be seen as imposing and impolite, while, as Tannen's dinner party account shows, less-active speakers could feel excluded in an imbalanced interaction.

Some other work has attended to speakers' perceptions of interactional patterns beyond the conversation-level balance of participation. For instance, some studies have tied fast-paced back-and-forths, or a preponderance of interruptions to social connection or "clicking" [53, 90]. Others have examined regimes of talk indicative of storytelling [2, 19, 38, 40, 51] or self-disclosure practices [82], suggesting emotionally and socially beneficial effects (e.g., [82]; see also [45]). Our approach provides a way of investigating similar questions and extending such analyses to more fine-grained models of talk-time sharing. We illustrate this in our empirical case study, detailed in Sections 4 and 5.1.

Talk-time considerations in computer-mediated settings. A broad range of research on computer-mediated interaction has focused on conversational dynamics in such settings. Our work is most directly relevant to studies of *synchronous* communication, including face-to-face interactions, audio calls, and video chat: the real-time nature of such settings renders the question of how participants share talk-time especially pertinent. However, as we discuss in Section 6, the distribution of talk remains pertinent for other channels such as instant messaging.

Several past studies have proposed ways of improving interactions (e.g., [76]). Some of these approaches have focused on mediating synchronous conversations, proposing interfaces or automated tools aimed at making speakers aware of their conversational practices (e.g., [48]). Of particular relevance to us, many of these interventions are concerned, at least in part, with whether the relative degree to which different people participate is balanced or skewed [16, 43, 48]. Other studies have documented gender-based imbalances in contributions to collaborative online efforts like Wikipedia [22, 47], in the service of making participation in such projects more equitable. These studies have mostly focused on analyzing and modulating overall levels of imbalance in interactions; our work surfaces additional dimensions—relating to the dynamics of talk—that designers and maintainers of such systems could consider.

A large body of work has also focused on designing conversational agents that can interact with people in realistic ways; such a project has gained prominence with recent developments in generative AI. Some scholars have drawn on accounts of conversational practices to examine chatbot behaviors, suggesting ways that these agents could be improved, or specifying new criteria for their evaluation [15, 64, 86]. Our framework, in providing an additional perspective on conversational dynamics, adds to these efforts.

3 Main Conversational Setting

To concretize the discussion of our approach, we start by describing the primary dataset we apply our framework to. We examine CANDOR, a large corpus of video chat dialogs [63], in which speakers were paired with people they'd never met, and simply told to have a conversation with no further stipulations. In using the framework to analyze this data, we will illustrate the types of conversational dynamics it can capture; to make sense of the characterizations our method produces, we will draw on survey responses from the speakers in these conversations.

The distribution of talk-time is relevant in any setting where multiple people interact, but for the purpose of an *initial* exploration, the CANDOR collection has several advantages. First, there are no pre-determined speaker roles. In more asymmetric scenarios—consider teacher-student dialogs—there would be additional constraints or norms governing how much talk each person takes up [11, 34]. No such a-priori expectations are applicable here, so we can more fully explore the possible space of talk-time sharing dynamics. Second, the dialogs are synchronous. In asynchronous settings, the distribution of talk-time would be impacted not just by the contingencies of real-time interaction, but also by absences as people close chat windows or email clients (e.g., [91]). This would introduce extra types of complexities when interpreting what the conversational dynamics mean. Finally, the conversations in the CANDOR corpus are dyadic. As such, our framework needs only to make a single distinction between the relative contributions of the two speakers, rather than deal with more complex configurations [80, 84], such as subgroups having side chats.

While analyzing talk-time sharing dynamics in dyadic video-chat conversations is our main focus in this work, we also test how our framework can be used in a different scenario. In Section 5.2, we discuss its application to a face-to-face, multi-speaker setting with asymmetric relations: the Supreme Court Oral Arguments. We further discuss factors that limit the generalizability of our framework and suggest avenues for addressing them in Section 6.

Data description. The CANDOR corpus consists of 1,656 conversations involving 1,456 unique speakers, and was collected by other researchers to facilitate large-scale studies of naturalistic conversation. The speakers, all located in the United States, were randomly paired with each other and instructed to have a 25-minute conversation, after which they'd be compensated. Both the audio and video of the conversation were recorded, and transcriptions were automatically produced using the AWS Transcribe API.
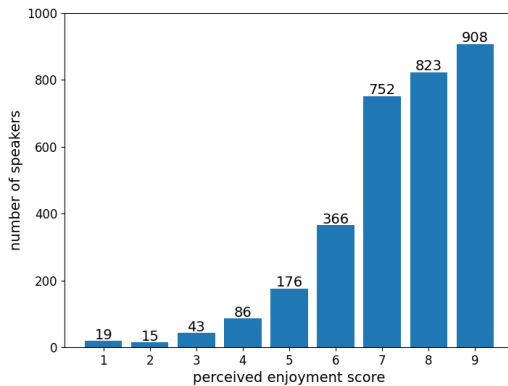

Fig. 1. Distribution of enjoyment rating among all speakers from the CANDOR corpus. Most speakers rate their conversations very highly, with 28.5% giving the maximum rating.

Table 1. Example post-conversation comments for a conversation in the CANDOR corpus (id: 36f338b1-bb0e-4c79-b40d-1c584913f262). Speaker A reported an enjoyment score of 9—the maximum value, while Speaker B reported an enjoyment score of 5.

	Example post-conversation comments
Speaker A	**Positive:** Everything went well, my partner was very pleasant, sociable, positive, warm, I actually learnt some new things during our conversation and I felt like I wanted to extend our conversation for a longer time and I think the feeling was mutual. I do believe my partner enjoyed the conversation too. **Negative:** There's nothing that want wrong during the conversation with my partner. My partner and I had a great time and we really enjoyed our conversation, we were both happy and actually it looked to me that if I didn't mention about the fact that our time was up my partner would have continued the conversation which means the conversation was really enjoyable and natural.
Speaker B	**Positive:** [NAME] was very conversational and had a lot of ideas for things to chat about. We had some things in common that allowed us to have deeper conversation than surface level. She was good at asking questions and allowing space for the answers, and appeared interested in what I was saying. **Negative:** [NAME] made space for me to talk and asked a lot of questions, but she spoke very quickly and didn't make much time for me to ask her questions. She also didn't seem to listen to my answers very specifically. She didn't allow much of a two way street in the conversation.

Speakers also completed surveys about their conversational experiences. We focus on two survey responses in particular. Speakers provided a numerical score of **how enjoyable** they perceived the conversation to be. As Figure 1 shows, these scores have a heavy positive skew. As such, when measuring the reported enjoyment over a set of conversations, we use the percentage of conversations that obtained the maximum score of 9, in addition to taking the mean enjoyment score. Speakers also wrote comments in which they separately discussed the **positive and negative aspects** of the interaction; examples are provided in Table 1.

We use these responses in complementary ways. If our framework makes a distinction between two sets of conversations with differing conversational dynamics, then differences in enjoyment score between these two sets would indicate that the dynamics we model have some bearing on speakers' experience—they contribute to different impressions of what happened, or co-occur with other interactional phenomena that produce different levels of enjoyment. The written comments can then provide us with additional clues about the types of experiences these dynamics point to. While we find these surveys extremely rich, we note that other paradigms—notably conversation analysis—would draw on what people said in the conversation, rather than their post-hoc assessments [35, 79] (see also [83]); we discuss how our computational approach could be extended to do this in Section 6.

For the sake of examining conversations that are comparable to each other, we exclude conversations where speakers reported significant technical issues like dropped calls, or where speakers didn't complete the study, i.e., who had particularly short conversations, or who didn't complete the survey. This results in 1,594 conversations for our analysis.

Measuring talk-time. Our framework starts from quantifying the amount of time each speaker spends talking. Transcripts in the CANDOR corpus consist of utterances from speakers and their start and end timestamps; we take the duration of an utterance to be its end time minus its start time, and the talk-time of a speaker to be the duration of all their utterances. Since the transcripts

are generated by an automated transcription system, which may introduce errors, we validate the speaking times derived from the transcripts against those obtained by directly processing the audio files.[1] We note that audio provides an alternative, perhaps richer, way to determine talk-time (e.g., [33]), which future work could explore.

4 A Framework for Analyzing the Sharing of Talk-Time

Our framework for examining the distribution of talk-time operates at two scales. We start with a straightforward, conversation-level measurement: to what extent is there an *imbalance* in the time taken by speakers? At a lower-level, we then consider how this imbalance dynamically changes as the conversation progresses. We structure the ensuing discussion accordingly: first, we introduce a simple measure of imbalance and explore how it relates to speakers' impressions of the conversation; then, we introduce a procedure to track and systematically analyze the dynamics of talk-time sharing, and identify different types of interactional patterns that—while reflective of equal levels of conversation-level imbalance—lead to diverging perceptions of the conversation. We make our framework's implementation publicly available as part of ConvoKit [7], enabling others to extend it and adapt it to other conversational settings. [2]

4.1 Preliminary measure: conversation-level imbalance

Our framework builds off a simple characterization of talk-time sharing: a conversation is balanced when all speakers talk for a similar amount of time and is imbalanced when some speakers take up more time than others. In what follows, we'll describe our framework for two-person conversations. Formally, we measure *imbalance* as the fraction of total talk-time in a conversation that's taken up by the most talkative speaker. We refer to the more talkative speaker as the *primary speaker* and the less talkative speaker as the *secondary speaker*.
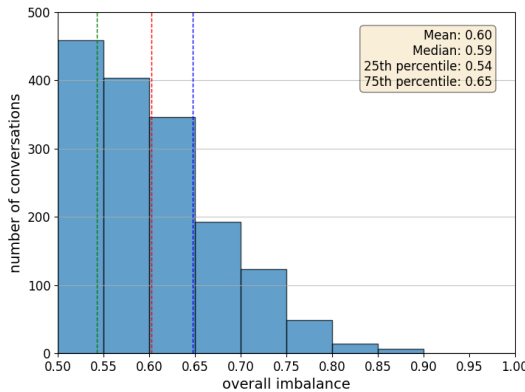
Fig. 2. Distribution of conversation-level imbalance in the CANDOR corpus.

Figure 2 shows the distribution of imbalance values over the CANDOR corpus. We note that many conversations are fairly balanced between the two speakers, (imbalance $\approx \frac{1}{2}$). However, for a substantial portion of the data, one speaker talks over twice as much as the other (imbalance $\geq \frac{2}{3}$).

---

[1]We also re-run all the analyses starting directly from the audio signal, and obtain qualitatively similar results. Details of these comparisons are included in Appendix A.3.

[2]https://convokit.cornell.edu/

Table 2. Phrases that most distinguish between balanced and imbalanced conversations, and representative excerpts including them. The comparisons are done separately for comments discussing positive and negative aspects. Additional examples are included in Table 7 and 8 in the Appendix.

	top distinctive phrases	example excerpts
Positive comments for **highly-balanced** conversations	both, lot in common, were able, were able to, like we, lot in, about our, able, able to, we both	• **We both** had a **lot in common**, so it made talking with each other super easy, and almost **like we** knew each other. • I feel **like we** connected and that makes it easier. • I felt open and **able to** share what is going on in my city and life.
Positive comments for **highly-imbalanced** conversations	chat, he, partner, talked lot, listener, else, im, is, she, bring	• I feel like my **chat partner** was great at asking me questions and keeping me talking. • **She talked a lot** and was very pleasent and smiled a lot. • Fortunately, I am a good **listener**.
Negative comments for **highly-balanced** conversations	our, our conversation, think of, of anything that, we had, anything that, both, her and, think the conversation, each other	• We flowed easily in **our conversation** and I am typing to get to fifty word count. • I think the rough parts were mainly how we tried to **think of** how to go back and forth, at first. • We **both** tend to be more introverted as we mentioned in the conversation.
Negative comments for **highly-imbalanced** conversations	than, more than, much, talked, me, too much, talked too, talked too much, he, ended up, ended	• **He talked** a lot **more than** I did. • I probably **talked too much**, but I felt like she wasn't giving very much in return so I don't think I had a choice. • **He ended up** opening up toward the end and actually provided me with some information that I might find useful for everyday life as well.

Speakers' perception. While basic, we see that our imbalance measure points to diverging conversational experiences; perhaps unsurprisingly (e.g., [29, 30]), people generally prefer balanced interactions over imbalanced ones. Comparing between *highly-balanced* conversations (top quartile of imbalance score) and *highly-imbalanced* conversations (bottom quartile) we see that in highly-balanced conversations 31.8% of the speakers give a maximum enjoyment score, versus only 25.3% in highly-imbalanced conversations. The respective mean enjoyment scores given by speakers in the two types of conversations are 7.53 and 7.25, respectively; these differences are statistically significant according to a Mann-Whitney U test ($p < 0.001$).

To further make sense of this distinction, we examine speakers' comments: what sorts of remarks are more likely to be associated with balanced versus imbalanced conversations? To compare comments across different parts of the data, we use the "Fightin' Words" statistical approach from Monroe et al. [54] that quantifies the extent to which a phrase (up to three words) occurs more frequently in comments from one part of the data (e.g., highly-balanced conversations) versus the other (e.g., highly-imbalanced conversations), relative to a word-frequency distribution bias.[3] This results in a list of phrases that are particularly distinctive in one versus the other part of the data (in our case, in highly-balanced versus highly-imbalanced conversations).

---

[3]Comparing to the prior—here, an uninformative Dirichlet prior—accounts for statistical irregularities stemming from overall word frequencies.

Table 2 contains distinguishing phrases for highly-balanced and highly-imbalanced conversations, along with comment excerpts that contain them (additional examples in Table 7 and 8 in the Appendix). For a more interpretable analysis, we use the Fightin' Words method to compare positive comments and negative comments separately. We see that speakers from highly-balanced conversations are more likely to refer to the joint aspects of the conversation: the most distinguishing words include "both", "we both", "our", "each other", and "lot in common", perhaps suggesting connections between balance and other forms of discursive convergence [30]. On the other hand, speakers from highly-imbalanced conversations are more likely to refer to individuals ("he", "she", "me", "i'm", and "partner"), or to the skewed distribution of talk ("talked too much"). We find this joint-vs-individualistic distinction in both the positive and negative comments.
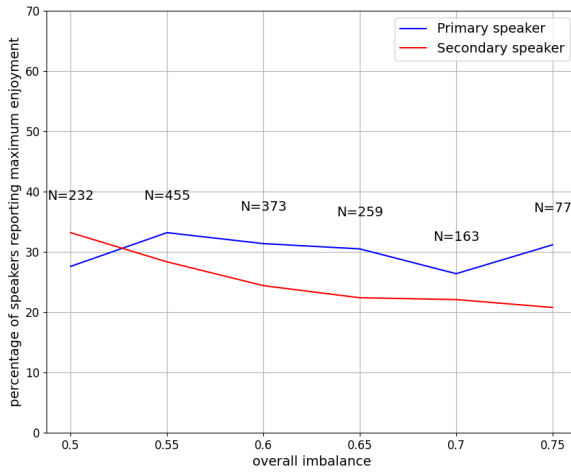

Fig. 3. Percentage of primary and secondary speakers reporting maximum enjoyment scores for conversations with different levels of imbalance. The number of conversations for each level is indicated as N. Levels with 30 or less conversations ($N < 30$) are not shown.

Difference between primary and secondary speakers. These patterns confirm that speakers often take notice imbalance of talk-time, and take actions or form judgements that reflect this characteristic of the conversation. However, do both speakers react to imbalance similarly? This question is especially pertinent in highly-imbalanced conversations where the roles of primary and secondary speaker become more salient. Figure 3 shows the relation between perceived enjoyment and level of imbalance for each type of speaker. We notice that the more the conversation is imbalanced, the less likely it is for the secondary speaker to enjoy it. This leads to an increasing gap between the way the primary and secondary speakers perceive their conversation.

This gap is also reflected when comparing the post-conversation comments of primary and secondary speaker in highly-imbalanced conversations, again using the Fightin' Words method (Table 3, extended in Table 9 and 10 in the Appendix).

Here, we find intuitive distinctions in the primary and secondary speakers' characterizations of how much their conversation partners talked. Additionally, we reveal a dichotomy on how talkativeness (or lack thereof) is perceived. Secondary speakers can perceive the talkativeness of their partners positively, appreciating that they "had [a] lot of" interesting "stories" to "share". But

Table 3. Phrases that distinguish most between comments by primary speakers and secondary speakers for highly-imbalanced conversations, and representative excerpts including them. The comparisons are done separately for positive comments and for negative comments. Additional examples are included in Table 9 and 10 in the Appendix.

	top distinctive phrases	example excerpts
Positive comments from **primary speaker**	questions, asked, but, probably, us, about my, agreed, some, young, good listener, each, interested	•She was a **good listener** and **asked** pertinent **questions** to get me talking. • I probably did the majority of the talking, but they **asked** good **questions** and gave interesting answers when I **asked** them **questions**. • She was also fun to talk to because she's a **good listener** and was understanding toward my situation.
Positive comments from **secondary speaker**	had lot, something, lot, stories, had similar, had lot of, share, friendly and, that my, life	•He **had a lot** to talk about so I let him tell me about those things for as long as he wanted to. • He had great life experience and **stories** that he **shared**. • My partner was very chatty and willing to **share** her views, while I was quite willing to listen.
Negative comments from **primary speaker**	her, little, too much, we had, flow, talked too, should, she was, she is, shy	• I think **she was** a **little** quiet by nature which is ok. • I think I probably **talked too much**. • I felt I led the conversation **too much** and perhaps dominated.
Negative comments from **secondary speaker**	different, hard, say, went, the conversation was, talked lot, life, honestly, didnt, kept, it was hard	• He talked way too much; **it was hard** to get a word in edgewise. • I did not know what to **say**. • My partner **talked a lot** about herself and **didn't** ask many questions, which made it harder for me to share as well.

they can also perceive it as detrimental to the conversation flow: "it was hard" to contribute when the other "talked [a] lot". This dichotomy is mirrored in the perceptions of the primary speakers: they can either feel connected when their partner is a "good listener" who shows "interest" by "asking questions", or they can regret "talk[ing] too much" and perceive the other person as being too "shy".

Such divergences in the perception of talk-time imbalance can result from different mechanisms, as prior work on floor-taking and conversational control suggests [5, 18, 55, 58]. A speaker can talk more because they have a lot to say or because their interlocutor isn't contributing; similarly a speaker can talk less because they don't have opportunities to do so or because they don't take those opportunities when they arise. Complementing our talk-time framework with analyses of utterance content could help disentangle the mechanisms at play, and in turn motivate design strategies for improving conversational dynamics.

4.2 Main framework: talk-time sharing dynamics

Conversation-level imbalance is a consequence of continuous negotiations between speakers at every point of the interaction. At any moment, a speaker might switch from taking the lead to focusing on the other speaker, or engage in a back-and-forth, sharing time more equally with their partner. Here, we introduce a method to capture and systematically examine such dynamics.

Following our approach, we end up with a *structured space* of talk-time sharing dynamics that can offer new insights into conversational data.

One basic element of our formalism is a *conversation window*: an arbitrarily small span of the conversation covering a specific length of time. We take a sliding window approach to capture the way the conversation progresses: each conversation is composed of multiple windows of K seconds, sliding in increments of L seconds (and thus overlapping for $K - L$ seconds), where $L < K$.[4] For each such conversation window, we measure the imbalance of talk-time, i.e., the fraction of talk-time taken up by the most talkative speaker *within that window* (as opposed to across the entire conversation, as in Section 4.1). This simple approach allows us to capture the dynamics of talk-time sharing: at what moments is talk-time dominated by one speaker versus the other, at what moments is talk-time equally distributed, and in what sequence do these different regimes occur?
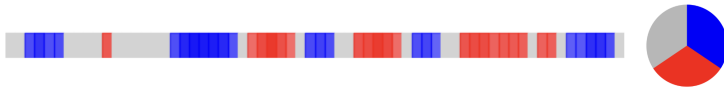

Fig. 4. Visualization of window-by-window talk-time sharing dynamics for an individual conversation (CANDOR ID: a7b14ca1-0b36-42b9-ad4a-50f0eb094035). The conversation level imbalance for this conversation is 0.51.

To aid our explanation, we introduce a way to visualize the sequence of talk-time sharing regimes in a given conversation (following the example in Figure 4). We use three colors to indicate the three respective regimes for a window: blue for windows where the (eventual conversation-level) primary speaker talks the most, red for windows where the (conversation-level) secondary speaker talks the most, and gray for windows in which the two speakers talk a similar amount of time (i.e., neither speaker's talk time exceeds $M\%$ of the total window talk-time; in our analysis on the CANDOR corpus we set $M = 60\%$).[5] The shades of blue and red correspond to the percentage of the window talk-time the respective speaker talks, with darker shades indicating higher percentages. The color-coded windows are illustrated in succession, without overlap between subsequent windows for clarity (even though their actual time overlaps). Finally, a pie-chart offers an overview by showing the percentage of blue, red, and gray windows; note that by construction, the percentage of blue windows will most often always be larger than that of red windows, because by construction blue corresponds to the speaker who talk the most overall.

This representation, as illustrated in Figure 4, reveals some aspects of the conversation that go beyond the conversation-level view offered by the simple measure of conversation-level imbalance described earlier. First, the overall primary speaker doesn't necessarily dominate the entire conversation, as shown in the presence of gray and red windows alongside the blue ones. Second, we see how the conversation progresses through multiple regimes of talk-time sharing: after some initial greetings and introductions (first gray window), the primary speaker takes the lead for most of the first half of the conversation (mostly blue and gray windows), narrating their mental health struggles during the initial months of the COVID-19 pandemic; the conversation then moves to a stretch led by the secondary speaker, who takes up the discussion of mental health to talk about their childrens' difficulties with school (long stretches of red windows).

---

[4]For our analysis, we use $K = 150s$ and $L = 30s$, arriving at these parameters after initial exploration of the CANDOR data; minor variations of K and L lead to similar results.

[5]We note that such distinctions are similar to those made in Edelsky [18] between singly- and jointly-developed floors, and in Eggins and Slade [19] and Gilmartin et al. [23] between "chunk" and "chat" sequences.

A structured space of talk-time sharing dynamics. Together, the regimes depicted in Figure 4 amount to a fairly balanced conversation (overall conversation-level imbalance = 0.51). However, in attending to the local dynamics of talk, we note that a different composition of regimes could have also resulted in the same conversation-level imbalance. For instance, a different, similarly-balanced conversation could consist of speakers who engage in more back-and-forths (i.e., more gray windows), rather than switching between longer stretches where one person leads.

We now propose a way of systematically making such distinctions, by mapping these representations to a structured space. We start by identifying three extreme cases: the primary speaker can **dominate throughout** a conversation, visualized as mostly blue windows; both speakers can solely engage in balanced **back-and-forth** exchanges, visualized as mostly gray windows; or, the secondary speaker can dominate in as many windows as possible—in such conversations, the primary and secondary speaker **alternate in dominating** the conversation, visualized as a mixture of blue and red windows (recall that by construction, the primary speaker almost always dominates more windows than the secondary, placing an upper limit on the possible proportion of red windows).

As depicted in Figure 5, these three conversational "stereotypes" bound a space of possible conversational dynamics. Of course, most conversations do not resemble the extreme cases, but by moving between these points, we can systematically account for richer interactional patterns. For instance, alternating periods of talk-time dominance can be interrupted by more equally-distributed back-and-forths; such dynamics correspond to conversations that lie on the axis between the "most gray" and "most red" extremes.
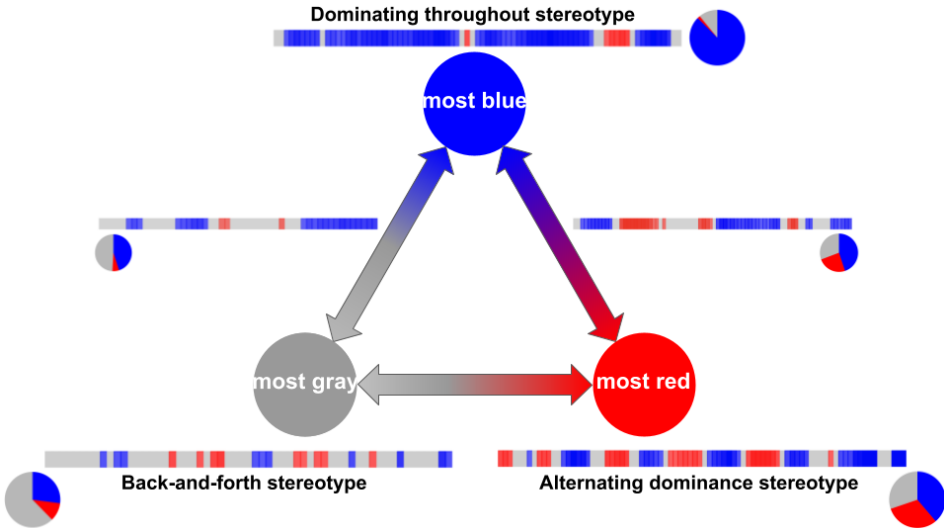

Fig. 5. A structured space of talk-time dynamics. The extremes are represented by blue, red and gray circles, and are accompanied by visualizations of example stereotypical conversations. Arrows indicate axes along which conversations can vary between each stereotype, and two example conversations that fall along these axes in between the stereotypes are included. Figure 6 shows the difference in reported enjoyment for conversations that lie at the extremes of the axes and Figure 7 shows how reported enjoyment varies along the bottom axis; variation along the other two axes are illustrated in the Appendix.

Speakers' perceptions. Starting from the three conversational stereotypes, we organize and systematically investigate the space of talk-time sharing dynamics. In particular, we note that

while the top of the space in Figure 5 (most blue) corresponds to very imbalanced conversations, the bottom (most gray, most red, and the space in-between) correspond to different ways of achieving balance in a conversation. Therefore, we can ask whether the way talk-time balance is achieved—either through back-and-forths or through alternating dominance—lead to different speaker perceptions, and thus refining our previous analysis that grouped all balanced conversations together.
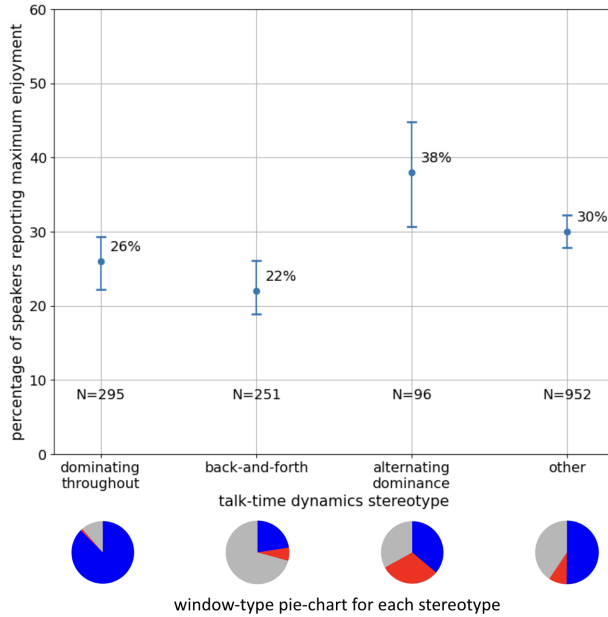

Fig. 6. Comparing the reported enjoyment of the three stereotypical talk-time dynamics: dominating throughout, back-and-forth, and alternating dominance; for reference, we include conversations that do not fit in any of these stereotypes (other). We show the percentage of maximum enjoyment scores and 95% confidence intervals obtained via bootstrap resampling. Pie charts are included to show the average percentages of window regime types for each stereotype.

Figure 6 shows the difference in perceived enjoyment for conversations that lie at the extremes of the axes in the CANDOR corpus. We see that even though stereotypical back-and-forth conversations (more than 60% gray windows)[6] and stereotypical alternating dominance conversations (more than 25% red windows) have relatively similar conversation-level imbalance (0.54 and 0.53 average conversation-level imbalance, respectively), there is a strong preference for the alternating dominance dynamics (average enjoyment score of 7.79 vs. 7.26 for back-and-forth, $p < 0.0001$ according to a Mann Whitney U test).[7]

In fact, back-and-forth conversations have very similar levels of enjoyment as the much more imbalanced dominated throughout conversations, which we take to be those with more than 75%

---

[6]We define boundary percentages heuristically, based on data distributions.

[7]To check that the observed difference in enjoyment is not explained by the difference in conversation-level imbalance, we also make a controlled comparison between back-and-forth and alternating dominance stereotypical conversations that are paired by their conversation-level imbalance (with tolerance of 0.005). We obtain 93 pairs which show no difference in conversation-level balance (Wilcoxon signed-rank test p-value = 0.42). In this tightly controlled set, we still observe a substantial difference in perceived enjoyment ($p < 0.05$).

blue windows (average conversation-level imbalance of 0.72; average enjoyment score of 7.2; Mann Whitney U test gives $p = 0.86$ comparing mean enjoyment between back-and-forth and dominating throughout conversations).

To further interpret the distinction between back-and-forth and alternating dominance conversations, we examine distinguishing phrases in the post-conversation comments, shown in Table 4 and Table 11 and 12 in the Appendix. Comparing positive comments, we note a stark difference in experience: while speakers in back-and-forth conversations are often only able to "keep the conversation going" by finding some common "topics", those engaging in alternating dominance were more likely to be interested in "listen[ing]" to "the other's" "stories" and "life experiences". These two distinct ways of connecting with a stranger also came with their specific challenges. Many of the negative comments for back-and-forth conversations pointed to a necessary effortfulness: participants "had to" do something to "keep" the conversation "flow" and avoid awkward "moments". The negative comments for alternating dominance were more diverse, though many questioned how genuine the interlocutor's interest actually was. In summary, the differences in the talk-time sharing dynamics of balanced conversations pointed to distinctions between familiar modes of interaction—in this setting, sharing extended stories or self-disclosures (e.g., [38, 51, 82]) versus (perhaps awkwardly) searching for conversation topics [41].

To extend our study beyond stereotypical conversations, we consider the variation across the gray-red axis. Here, we examine conversations with varying proportions of red windows. To separate this analysis from our preceding discussion of conversation-level imbalance, we analyze conversations where the percentage of blue windows is below 50% (the "bottom" of the space illustrated in Figure 5). Figure 7 shows that reported enjoyment increases as the proportion of red windows increases—i.e., the secondary speaker leads for longer periods of time.[8]

Changes in dominant speaker. We explore one further (even finer-grained) type of variation that our framework surfaces: we differentiate between conversations with alternating dominance dynamics based on the number of "alternations," i.e., the number of times the dominant speaker role *flips* between the primary and secondary speaker. At one extreme, a conversation could contain only one such flip, i.e., a sequence of blue windows followed by a sequence of red windows. At the other extreme, there could be many such flips; for instance, the conversation depicted in Figure 4 has 8 flips.[9] Even if they have the same overall levels of participation, these two scenarios seem different, at least at an intuitive level: single-flip conversations seem more akin to two monologues stitched together, while speakers trade off the leading role in multi-flip scenarios, making the conversation potentially more dialogic or reciprocal. In order to examine flips while controlling for the relative proportion of different regimes (i.e., as depicted in the pie charts), we match each single-flip conversation with one that has 3 or more flips, such that they have the same proportion of blue, red, and gray windows (with a tolerance of 2%). This matching process leads to a small number ($N = 90$) of tightly controlled pairs. In 53.3% of these pairs, the multi-flip conversations receive higher aggregate enjoyment scores,[10] compared to only 37.8% of pairs where the single-flip conversations receive higher scores (in the rest of the cases the scores are equal); $p = 0.06$ according to a one-tailed sign test.

---

[8]Variation in enjoyment scores along the other two axes are depicted in the Appendix: variation across the blue- red axis yields differences in speaker perception that match with the decrease in conversation-level imbalance, variation across the blue- gray axis however does not show any significant differences in speaker perception, in spite of also corresponding to a decrease in conversation-level imbalance. Comparing these two ways of going from imbalanced conversations to balanced conversations further justifies accounting for finer-grained dynamics.

[9]Different stretches of blue or red windows could be punctuated by gray windows throughout. For this analysis we ignore all gray windows and compute blue-red or red-blue transitions after removing gray windows.

[10]To obtain a conversation-level aggregate enjoyment scores, we add the enjoyment scores from each speaker.

Table 4. Most distinguishing phrases between comments for stereotypical alternating dominance and back-and-forth conversations, and representative excerpts including them. The comparisons are done separately for positive comments and for negative comments. Additional examples are included in Table 11 and 12 in the Appendix.

	top distinctive phrases	example excerpts
Positive comments for **alternating dominance** conversations	stories, to listen, listen, stranger, life experiences, the other, myself, his, good conversation, super	• We both had interesting **stories** to tell, so that was also very enjoyable. • I feel like we were both able **to listen** actively to one another, and bring curiosity to the conversation. • We both listened when **the other** was speaking and asked follow-up questions.
Positive comments for **back-and-forth** conversations	which, the conversation going, conversation going, topics, keep the, keep the conversation, games, going, well was, discussed	• We both took turns talking and **keeping the conversation going** with each other. • The flow of the conversation went well in which we kept bringing up **topics** to speak on and it went smoothly. • Not a lot outside of talking about **games**.
Negative comments for **alternating dominance** conversations	go, one, interested, she, into the, minutes, get, wasnt really, through, few minutes	• My partner didn't seem **interested** in me at all. • As soon as we were a **few minutes** over, **she** wanted to end the conversation. • **She** laughed at stuff that **wasn't really** funny several times, so I couldn't tell if it was for humor or nerves.
Negative comments for **back-and-forth** conversations	some, going, im, most, had to, went well, understand, flow, kept, topics, moments	• The conversation had **some** awkward **moments** and silent pauses. • He didn't have a lot of talking points and I felt as though I **had to** keep asking questions to keep the conversation **going**. • The conversation **flow** seemed forced especially when one topic had been discussed and we need to find another.

Mixed dynamics. Conversations can also transition between different talk-time sharing regimes: for instance, the conversation visualized in Figure 8 starts with a back-and-forth dynamic before transitioning to an alternating dominance regime. Accounting for this heterogeneity is straightforward: we apply our framework to different parts of the conversation. For this analysis, we focus on capturing when conversations switch from one regime to another partway through,[11] by using our framework to characterize the first 60% and the last 60% of each conversation. [12] We find that the most common transitions in this setting involved going from a back-and-forth to alternating dominance regime, and the other way around (13 conversations captured for each transition direction, additional examples in Figure 21 and Figure 22), while mixes involving the dominating-throughout regime are relatively rare (Figures 23 and 24).

---

[11]We note that this analysis can easily be extended to account for mixtures of more than two regimes.
[12]We use 60%, allowing overlap between the two halves of the conversation, rather than a strict 50% cut, to account for the transition in between, enhancing robustness in capturing gradual shifts.
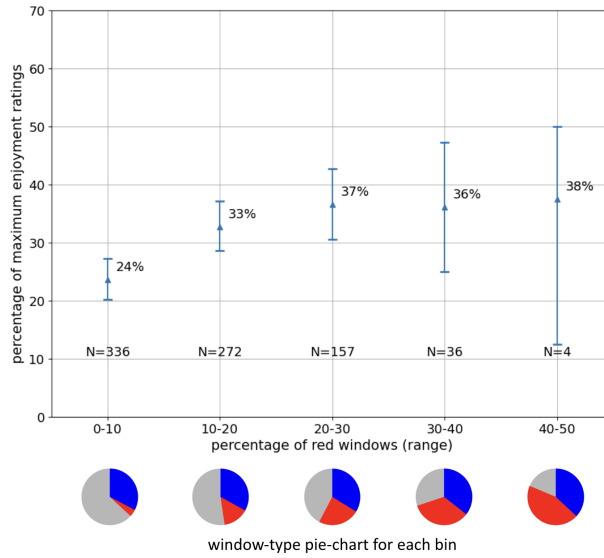
Fig. 7. After controlling for the proportion of blue windows, we compare reported enjoyment for varying proportions of red windows.

Fig. 8. Visualization of a conversation with talk-time sharing dynamics stereotype transitioning from back-and-forth to alternating dominance (CANDOR ID: 6fb27106-8840-406e-a800-99d1516191ce). Additional examples of conversations with mixed regimes are included in Figures 20, 21, 22, 23, and 24 in the Appendix.

5 Case Studies

To further illustrate the utility of our framework, we present two case studies. First, we conduct an exploratory analysis on how speaker characteristics relate to talk-time sharing dynamics in the CANDOR corpus. Second, we apply our framework to a dataset of Supreme Court Oral Arguments—comprising structured, multi-speaker interactions—to demonstrate the framework's applicability to a contrasting setting. Together, these applications underscore the framework's extensibility and generalizability.

5.1 Exploratory analysis: Speaker characteristics and time sharing dynamics

Here, we illustrate how our framework can be used to examine the tendencies of speakers with different characteristics to engage in different talk-time sharing dynamics. Rather than trying to make broad claims about the relation between social identity and interaction style based on the CANDOR corpus, we present this analysis as an example of another potential application of the framework that can tie into existing scholarship.

Past work has suggested connections between speakers' characteristics—including gender and cultural background [5, 28, 87, 99, inter alia]—and conversational behaviors. Tannen [88], for instance, posits that the behaviors of guests at the dinner party she analyzes may reflect the contexts

in which they were socialized, and points to *conversational styles* that pervade across different interactions. In the CANDOR corpus, we see that speakers themselves link their conversational experiences to assessments of their own, or their partners', character traits (e.g., "introvert," "shy," "talkative"). Besides contributing to the existing literature on psychological and cultural determinants of conversational dynamics, an investigation of how talk-time sharing relates to speakers' attributes could also inform designers interested in promoting better, or more equitable conversations (see Section 6 for further discussion). We approach this in two ways: we examine whether speakers' roles are consistent across their conversations; we then analyze how these roles, and the interactional patterns they take part in, are related to demographic factors.

Role consistency across conversations. Do speakers' characteristics like "shyness" or "talkativeness" pervade across the conversations they take part in? We address this question in terms of whether people are consistent in being the primary or secondary speaker, making use of the subset of participants in the CANDOR study who take part in multiple conversations. For the 148 speakers who participate in multiple highly-imbalanced conversations, we randomly select two conversations and compare their role in them. We find that 88.1% of these speakers are either the primary speaker in both conversations, or the secondary speaker in both—a much higher rate than agreement by chance (Cohen's $\kappa = 0.76$). This suggests that in this setting, there is a strong tendency for speakers to stick to the same role across multiple conversations, even with different conversation partners.

Relation to demographic attributes. What speaker characteristics might relate to their conversational roles? We focus on gender and age, two attributes that speakers self-report in their survey responses.[13] Over the 781 conversations involving a male speaker paired with a female speaker, we see that in 56% of them, the primary speaker is male ($p < 0.001$ per a two-tailed signed test), consistent with several past studies demonstrating the propensity of male speakers to dominate interactions. Over the 1,295 conversations where speakers differed in age by at least 3 years, we see that in 65% of them, the primary speaker is the older of the pair ($p < 0.001$ per a two-tailed signed test).

As noted above, these analyses presented here are exploratory and are meant to demonstrate potential uses of our framework, rather than making general distinctions between demographics. Additional information and data would be needed to determine what accounts for these findings, and whether they are specific to the CANDOR setting, especially given the complex nature of social identity [77]—for instance, to what extent do differences in conversational dynamics reflect the speakers, versus their partners' *perceptions* of them [57, 81]? Nonetheless, this investigation provides examples of how our framework can be applied to start addressing such questions.

5.2 Application to Supreme Court Oral Arguments

While the main focus in this work is on the video-chat dataset, our framework can—with some adaptation—be used to examine a diversity of other settings. As a demonstration, we apply our framework to a collection of Supreme Court Oral Arguments [78]. This domain differs from our primary video-chat setting in several key ways, allowing us to illustrate the generalizability of our framework. In contrast to a casual, dyadic exchange involving two strangers on equal footing, there are more than two speakers occupying different institutional roles: lawyers argue for or against a case, while justices listen to these arguments and ask questions. This has clear implications for speakers' expected share of talk-time; lawyers almost always speak much more than justices. Finally, oral arguments are generally face-to-face as opposed to mediated by a video-chat platform.

---

[13]In terms of gender, speakers can report either "Male", "Female", "Other", or "prefer not to answer".

Data description. The Supreme Court Oral Arguments dataset [7, 13] contains transcripts of oral arguments for cases from the US Supreme Court dating back to 1955. Each argument involves nine Supreme Court justices and lawyers from two sides—representing the petitioner and the respondent, respectively—who take turns to present their arguments and answer the justices' questions. For brevity, here we only consider the petitioner's turn and discuss the results for the respondent's turn in the Appendix (Section A.4). We filter out cases with missing or additional roles from any party, resulting in 4,528 Supreme Court cases for our analysis. The dataset also includes professional transcripts with time-stamps, which we use to determine talk-time.

The data also includes the nine justices' votes for each case, which are cast after the oral arguments. For this demonstration, we examine whether the distribution of votes is related to the oral argument discussion dynamics. Here, we focus on comparing discussions that preceded a unanimous vote—where all justices align—with those that preceded a divided vote—which is split 5-to-4. In the data, there are 1,760 unanimous votes and 816 divided ones.

Adapting the framework. To use our framework in this setting, we first need to account for the fact that these interactions have more than two speakers. Here, we treat all the lawyers from one side as a single speaker, as well as all of the justices; as such, during the petitioners' turn, there are effectively two speakers: the lawyers and the justices. Future work could explore more elaborate adaptations to group conversations (see Section 6).

In adapting the framework, we also need to account for the asymmetric nature of the setting, since lawyers and justices occupy very different roles with different expectations for their conduct. Indeed, on average, lawyers take up more than twice as much talk-time as justices (median overall conversation imbalance of 0.74). To reflect the institutionalized nature of these speaker roles, we by default assign lawyers and justices to the primary and secondary speaker roles respectively (blue and red in our visualization). We additionally modify our criteria for determining which speaker dominates each conversation-window: lawyers need to talk more than M_L = 80% of the speaking time to be considered as dominating that window (blue in our visualization), while justices only need to talk more than M_J = 40% to be considered as dominating the window (red in our visualization). If neither of these thresholds is reached, none of the parties are considered to dominate the window (gray in the visualization). We also modify our criteria for the three stereotypical conversations from Figure 5: we take back-and-forth conversations to be those with more than 40% gray windows, alternating dominance conversations as those with than 40% red windows, and dominating throughout conversations as those with more than 70% blue windows.[14]

Case outcome and talk-time sharing dynamics. Figure 9 shows the proportion of cases won by the petitioner, split up by whether the justices' votes were unanimous or divided, and by the talk-time sharing dynamics of the oral argument. In cases where petitioners dominate the oral argument throughout, their share of wins is higher. Alternating dominance dynamics correspond to lower proportions of wins, perhaps reflecting more substantial scrutiny from the justices [36, 78]. These differences are particularly notable in divided cases, where perhaps more weight is placed on the petitioners' argumentation [10]. These findings suggest that the conversational dynamics that our framework surfaces connect to the discursive and judicial processes taking place in Supreme Court cases. More broadly, they provide a starting illustration of how our framework can be applied in new settings, with appropriate modifications.

---

[14]In addition, we adjust the sliding window parameters K = 120s and L = 30s to account for shorter justices' talk-time in local spans. Minor variations of these parameters lead to similar results.
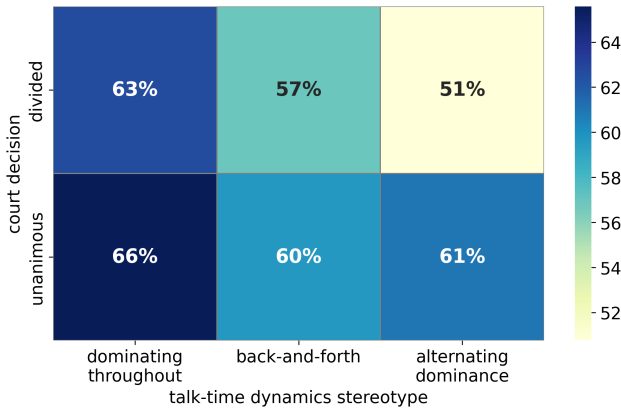
Fig. 9. Comparing the share of cases won by petitioners across different conversation stereotypes and court decision types.

6   Discussion

In this work, we introduce a framework to model the dynamic distribution of talk-time in a conversation. Starting from a basic account of conversation-level (im)balance between more or less talkative speakers, our framework goes on to make finer-grained distinctions between various lower-level dynamics. We derive a structured space of talk-time sharing dynamics bounded by three stereotypical conversation patterns—conversations where one speaker continuously dominates the talk-time, conversations where two speakers alternate in dominating the talk-time, and back-and-forth conversations where neither speaker fully takes the lead. We show, via a case study of a collection of video chats, how this perspective can enrich past accounts. In particular, while we replicate existing findings that speakers generally enjoy balanced conversations (e.g., [29]), we also find that even among fairly balanced conversations, speakers report very different experiences, depending on the particular dynamics of talk-time sharing involved.

Adapting to other modes of communication and communication settings. While our main focus is on a specific conversational scenario—dyadic, synchronous, and role-neutral—future work could explore how talk-time dynamics play out in different scenarios. This might require adapting some of the framework's basic elements, as demonstrated in our exploration of Supreme Court Arguments in Section 5.2. For example, while in dyadic conversations imbalance can straightforwardly be measured as a fraction of talk-time, this would not be suitable for group conversations involving more than two parties. While we used simple heuristics to treat the Supreme Court interactions as effectively dyadic, future work could consider more sophisticated approaches, such as entropy-based measures previously used to measure balance in group conversations [56]. In contexts where speakers have pre-determined roles that strongly condition talk-time sharing dynamics—such as in the Supreme Court Oral Arguments—those roles could replace or be added to the primary and secondary speaker distinction; this could help to more accurately model intuitive judgements of talk-time dominance, and also enable investigations of departures from expected dynamics. Finally, it would be impossible to measure talk-time based on the time duration of utterances in written interactions; simple adaptations could instead use the number of words (or even drafting-time) of an utterance instead. More generally, such simple extensions could expand the impact of the proposed framework, and serve as an additional basis for comparing interactions across different communication media [3, 74].

The structured space of talk-time sharing dynamics could be used to compare across different contexts. For example, one could also add to accounts of cultural variation in conversational practices [12, 15, 75, 87], going beyond the American setting considered here. One could also examine the relation to social relationships between participants. For instance, we find a dispreference for back-and-forths; this is somewhat surprising given past work linking such exchanges to feeling connected to one's conversational partner [90]. We speculate that in other settings, where conversational partners are closer to each other, such dynamics might point to flowing banter facilitated by shared understandings (e.g., [94]); in our setting, which involves interactions between strangers, such a dynamic instead seems to point to an awkward search for conversation topics (as indicated by speakers' own accounts).

Examining linguistic processes. Our approach could be enriched by accounting for the language used within conversations. While our study used post-hoc reports from speakers to make sense of the various talk-time sharing dynamics we surfaced, future work could examine what was said within and across different regimes of talk. Here, we could draw on Gilmartin et al. [23], who proposes an ontology to label different segments of talk (e.g., as storytelling, or as gossiping), or on Nguyen et al. [55], who track the progression of topics in a conversation.

How do speakers' actions shape the talk-time sharing dynamics observed? Past work points to the various interactional moves that people can make in negotiating who has the floor, who should take the next turn, and when to switch [18, 39, 66, inter alia]. Our method could be paired with such finer-grained approaches: for instance, in tandem with locating conversations with alternating dominance patterns, we could examine the linguistic or interactional phenomena present in particular cases that might account for switches in lead speaker.

Design implications. The talk-time sharing framework has several potential applications relating to the design of computer mediated communication platforms, for both human-human and human-AI communication. First, it can be used as a diagnostic tool: platform maintainers could analyze interactions and their talk-time sharing dynamics, drawing connections to inequities or to particular goals like participants' enjoyment or discussion constructiveness. This could help inform ways of improving consequential processes like collaborative production [22, 31, 47], participatory governance [21, 61], or the provision of care [14].

Second, we suggest that designers who focused on computer-mediated communication settings should attend to the dynamics of talk-time sharing. A number of past studies have proposed interfaces or automated tools to improve interactions, in part by making participants aware of the relative amount they're contributing [16, 43, 48, 92]. Our findings illustrate that it is important to distinguish between types of interaction that appear to have the same conversation-level imbalance of talk, since different dynamics reflect diverging levels of speakers' enjoyment. Other work points to imbalances in participation related to factors like gender and age [22, 47], which can adversely impact collaborative projects. We surface ways that such inequities can relate to finer-grained interaction patterns as well.

Third, our framework could help to more extensively explore design spaces; for instance, interfaces could depict more complex interactional patterns such as the ones we've highlighted, while interventions aimed at mediating discussions could better account for the dynamics of talk to improve their efficacy while minimizing their disruptiveness. These ideas could enrich present approaches aimed at improving the treatment of newcomers to a platform [32], alleviating social anxiety [42], encouraging empathy toward victims of cyberbullying [89], diffusing tensions [8], increasing constructiveness in teams [6, 56], alongside many other avenues [76]. Future work could use our framework to design UI interventions that would inform speakers of ongoing talk-time

sharing dynamics—e.g., by displaying a version of our visualization. We note that such interventions depend on ways of more rigorously relating talk-time sharing dynamics to desired outcomes, requiring careful user studies or causal analyses (e.g., see [97]).

Finally, our work could inform the design and evaluation of conversational agents—an especially pertinent focus, given recent developments in AI framed around conversational interfaces, like ChatGPT. A range of existing work has pointed to ways that people expect AI agents to behave from a conversation behavior standpoint [98] and ways in which they fall short of realistic and fluent interaction [15, 64, 86], suggesting that careful attention to conversational mechanics can help with diagnosing existing issues while calibrating future expectations. Our framework can help with these efforts, in making the dynamic distribution of talk-time available for analysis. Additionally, in revealing and making distinctions within a space of possible dynamics, we add to accounts (e.g., [15]) that urge designers seeking to automate interaction to contend with the diversity of human conversation.

Acknowledgments

References

[1] Paul Baker, Costas Gabrielatos, Majid Khosravinik, Michał Krzyżanowski, Tony McEnery, and Ruth Wodak. 2008. A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & society* 19, 3 (2008), 273–306. Publisher: Sage Publications Sage UK: London, England.

[2] Janet B Bavelas, Linda Coates, and Trudy Johnson. 2000. Listeners as co-narrators. *Journal of personality and social psychology* 79, 6 (2000), 941. Publisher: American Psychological Association.

[3] Jonah Berger and Raghuram Iyengar. 2013. Communication Channels and Word of Mouth: How the Medium Shapes the Message. *Journal of Consumer Research* 40, 3 (Oct. 2013), 567–579. doi:10.1086/671345

[4] Frances R. Bilous and Robert M. Krauss. 1988. Dominance and accommodation in the conversational behaviours of same- and mixed-gender dyads. *Language & Communication* 8 (1988), 183 – 194. doi:10.1016/0271-5309(88)90016-X 00247.

[5] Victoria L. Brescoll. 2011. Who Takes the Floor and Why: Gender, Power, and Volubility in Organizations. *Administrative Science Quarterly* 56, 4 (Dec. 2011), 622–641. doi:10.1177/0001839212439994 Publisher: SAGE Publications Inc.

[6] Hancheng Cao, Vivian Yang, Victor Chen, Yu Jin Lee, Lydia Stone, N'godjigui Junior Diarrassouba, Mark E. Whiting, and Michael S. Bernstein. 2021. My Team Will Go On: Differentiating High and Low Viability Teams through Team Interaction. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3 (Jan. 2021), 230:1–230:27. doi:10.1145/3432929

[7] Jonathan P. Chang, Caleb Chiam, Liye Fu, Andrew Wang, Justine Zhang, and Cristian Danescu-Niculescu-Mizil. 2020. ConvoKit: A Toolkit for the Analysis of Conversations. In *Proceedings of SIGDIAL*. https://www.aclweb.org/anthology/2020.sigdial-1.8

[8] Jonathan P. Chang, Charlotte Schluger, and Cristian Danescu-Niculescu-Mizil. 2022. Thread With Caution: Proactively Helping Users Assess and Deescalate Tension in Their Online Discussions. In *Proceedings of CSCW*. 545:1–545:37. doi:10.1145/3555603

[9] Herbert Clark and Dale Schunk. 1980. Polite responses to polite requests. *Cognition* 8, 2 (1980), 111–143. doi:10.1016/0010-0277(80)90009-8

[10] Tom S. Clark, B. Pablo Montagnes, and Jörg L. Spenkuch. 2022. Politics from the Bench? Ideology and Strategic Voting in the U.S. Supreme Court. *Journal of Public Economics* 214 (Oct. 2022), 104726. doi:10.1016/j.jpubeco.2022.104726

[11] Steven E Clayman, Marc N Elliott, John Heritage, and Megan K Beckett. 2010. A watershed in White House journalism: Explaining the post-1968 rise of aggressive presidential news. *Political Communication* 27, 3 (2010), 229–247. Publisher: Taylor & Francis.

[12] Jenny Cook-Gumperz. 2006. *The Social Construction of Literacy*. Cambridge University Press. Google-Books-ID: 9v2LexyCRDsC.

[13] Cristian Danescu-Niculescu-Mizil, Lillian Lee, Bo Pang, and Jon Kleinberg. 2012. Echoes of Power: Language Effects and Power Differences in Social Interaction. In *Proceedings of WWW*. doi:10.1145/2187836.2187931

[14] Munmun De Choudhury, Sanket S. Sharma, Tomaz Logar, Wouter Eekhout, and René Clausen Nielsen. 2017. Gender and Cross-Cultural Differences in Social Media Disclosures of Mental Illness. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. Association for Computing Machinery, New York, NY, USA, 353–369. doi:10.1145/2998181.2998220

[15] Mark Dingemanse and Andreas Liesenfeld. 2022. From text to talk: Harnessing conversational corpora for humane and diversity-aware language technology. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. 5614–5633.

[16] Hyo Jin Do, Ha-Kyung Kong, Jaewook Lee, and Brian P. Bailey. 2022. How Should the Agent Communicate to the Group? Communication Strategies of a Conversational Agent in Group Chat Discussions. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (Nov. 2022), 1–23. doi:10.1145/3555112

[17] Kent Drummond. 1989. A backward glance at interruptions. *Western Journal of Speech Communication* 53, 2 (Aug. 1989), 150–166. doi:10.1080/10570318909374297 00072.

[18] Carole Edelsky. 1981. Who's Got the Floor? *Language in Society* 10, 3 (1981), 383–421. https://www.jstor.org/stable/4167262 Publisher: Cambridge University Press.

[19] Suzanne Eggins and Diana Slade. 2004. *Analysing Casual Conversation*. Equinox Publishing Ltd. Google-Books-ID: cTEkm6WzfNgC.

[20] Nicolas Fay, Simon Garrod, and Jean Carletta. 2000. Group Discussion as Interactive Dialogue or as Serial Monologue: The Influence of Group Size. *Psychological Science* (2000), 481–486.

[21] Seth Frey, P. M. Krafft, and Brian C. Keegan. 2019. "This Place Does What It Was Built For": Designing Digital Institutions for Participatory Change. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW (Nov. 2019), 32:1–32:31. doi:10.1145/3359134

[22] Jana Gallus and Sudeep Bhatia. 2020. Gender, power and emotions in the collaborative production of knowledge: A large-scale analysis of Wikipedia editor conversations. *Organizational Behavior and Human Decision Processes* 160 (Sept. 2020), 115–130. doi:10.1016/j.obhdp.2020.02.003

[23] Emer Gilmartin, Carl Vogel, and Nick Campbell. 2018. Chats and Chunks: Annotation and Analysis of Multiparty Long Casual Conversations. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Nicoletta Calzolari, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi, Koiti Hasida, Hitoshi Isahara, Bente Maegaard, Joseph Mariani, Hélène Mazo, Asuncion Moreno, Jan Odijk, Stelios Piperidis, and Takenobu Tokunaga (Eds.). European Language Resources Association (ELRA), Miyazaki, Japan. https://aclanthology.org/L18-1309

[24] Erving Goffman. 1981. *Forms of Talk*. University of Pennsylvania Press. Google-Books-ID: Z3bvx_T4Zu8C.

[25] Julia A. Goldberg. 1990. Interrupting the discourse on interruptions. *Journal of Pragmatics* 14, 6 (Dec. 1990), 883–903. doi:10.1016/0378-2166(90)90045-F 00217.

[26] Barbara J Grosz, Aravind K Joshi, and Scott Weinstein. 1994. Centering: a framework for modelling the coherence of discourse. (1994).

[27] Barbara J. Grosz and Candace L. Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics* 12 (July 1986), 175–204. 03990.

[28] John J. Gumperz. 1982. *Discourse Strategies*. Cambridge University Press. Google-Books-ID: aUJNgHWl_koC.

[29] Andrew J. Guydish, J. Trevor D'Arcey, and Jean E. Fox Tree. 2021. Reciprocity in Conversation. *Language and Speech* 64, 4 (Dec. 2021), 859–872. doi:10.1177/0023830920972742 Publisher: SAGE Publications Ltd.

[30] Andrew J. Guydish and Jean E. Fox Tree. 2023. In Pursuit of a Good Conversation: How Contribution Balance, Common Ground, and Conversational Closings Influence Conversation Assessment and Conversational Memory. *Discourse Processes* (Jan. 2023). https://www.tandfonline.com/doi/full/10.1080/0163853X.2022.2152552 Publisher: Routledge.

[31] Aaron Halfaker, R. Stuart Geiger, Jonathan T. Morgan, and John Riedl. 2013. The Rise and Decline of an Open Collaboration System. *American Behavioral Scientist* 57, 5 (May 2013), 664–688. doi:10.1177/0002764212469365

[32] Aaron Halfaker, Bryan Song, D. Alex Stuart, Aniket Kittur, and John Riedl. 2011. NICE: Social Translucence Through UI Intervention. 101–104. doi:10.1145/2038558.2038575

[33] Mattias Heldner and Jens Edlund. 2010. Pauses, gaps and overlaps in conversations. *Journal of Phonetics* 38, 4 (2010), 555–568. Publisher: Elsevier.

[34] John Heritage and Paul Drew. 1992. Talk at work. *Interaction in institutional settings* 2 (1992). Publisher: Cambridge University Press.

[35] Elliott M Hoey and Kobin H Kendrick. 2017. Conversation analysis. *Research methods in psycholinguistics and the neurobiology of language: A practical guide* (2017), 151–173. Publisher: Wiley Online Library.

[36] Tonja Jacobi and Dylan Schweers. 2017. Justice, interrupted: The effect of gender, ideology, and seniority at Supreme Court oral arguments. *Va. L. Rev.* 103 (2017), 1379. Publisher: HeinOnline.

[37] Deborah James and Janice Drakich. 1993. Understanding gender differences in amount of talk: A critical review of research. In *Gender and conversational interaction.* Oxford University Press, New York, NY, US, 281–312.

[38] Gail Jefferson. 1978. Sequential aspects of storytelling in conversation. In *Studies in the organization of conversational interaction.* Elsevier, 219–248.

[39] Gail Jefferson. 1993. Caveat speaker: Preliminary notes on recipient topic-shift implicature. *Research on Language and social interaction* 26, 1 (1993), 1–30. Publisher: Taylor & Francis.

[40] Gail Jefferson. 2015. *Talking about Troubles in Conversation.* Oxford University Press. Google-Books-ID: CntVD-wAAQBAJ.

[41] Dan Jurafsky, Rajesh Ranganath, and Dan McFarland. 2009. Extracting social meaning: Identifying interactional style in spoken conversation. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics.* 638–646.

[42] Matthew K. Miller, Martin Johannes Dechant, and Regan L. Mandryk. 2021. Meeting You, Seeing Me: The Role of Social Anxiety, Visual Feedback, and Interface Layout in a Get-to-Know-You Task via Video Chat.. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.* 1–14.

[43] Taemie Kim, Agnes Chang, Lindsey Holland, and Alex Sandy Pentland. 2008. Meeting mediator: enhancing group collaborationusing sociometric feedback. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work.* ACM, San Diego CA USA, 457–466. doi:10.1145/1460563.1460636

[44] Dennis Kurzon. 2007. Towards a typology of silence. *Journal of Pragmatics* 39, 10 (Oct. 2007), 1673–1688. doi:10.1016/j.pragma.2007.07.003

[45] William Labov, Richard Bauman, Jerome Bruner, Rom Harré, Alan D Hansen, Luke Moissinac, Cristal Renteria, Eliana Razo, Leor Cohen, Gabriella Modan, and others. 2010. *Telling stories: Language, narrative, and social life.* Georgetown University Press.

[46] Robin Lakoff. 1973. The logic of politeness: Or, minding your p's and q's. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, Vol. 9. Chicago Linguistic Society, 292–305. Issue: 1.

[47] Shyong (Tony) K. Lam, Anuradha Uduwage, Zhenhua Dong, Shilad Sen, David R. Musicant, Loren Terveen, and John Riedl. 2011. WP:clubhouse? an exploration of Wikipedia's gender imbalance. In *Proceedings of the 7th International Symposium on Wikis and Open Collaboration (WikiSym '11).* Association for Computing Machinery, New York, NY, USA, 1–10. doi:10.1145/2038558.2038560

[48] Gilly Leshed, Diego Perez, Jeffrey T. Hancock, Dan Cosley, Jeremy Birnholtz, Soyoung Lee, Poppy L. McLeod, and Geri Gay. 2009. Visualizing real-time language-based feedback on teamwork behavior in computer-mediated groups. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, Boston MA USA, 537–546. doi:10.1145/1518701.1518784

[49] Stephen C Levinson. 2016. Turn-taking in human communication–origins and implications for language processing. *Trends in cognitive sciences* 20, 1 (2016), 6–14. Publisher: Elsevier.

[50] Neil G MacLaren, Francis J Yammarino, Shelley D Dionne, Hiroki Sayama, Michael D Mumford, Shane Connelly, Robert W Martin, Tyler J Mulhearn, E Michelle Todd, Ankita Kulkarni, and others. 2020. Testing the babble hypothesis: Speaking time predicts leader emergence in small groups. *The Leadership Quarterly* 31, 5 (2020), 101409. Publisher: Elsevier.

[51] Jenny Mandelbaum. 2012. Storytelling in conversation. *The handbook of conversation analysis* (2012), 492–507. Publisher: Wiley Online Library.

[52] Marianne Schmid Mast. 2002. Dominance as Expressed and Inferred Through Speaking Time. *Human Communication Research* 28, 3 (2002), 420–450. doi:10.1111/j.1468-2958.2002.tb00814.x _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-2958.2002.tb00814.x.

[53] Daniel A. McFarland, Dan Jurafsky, and Craig Rawlings. 2013. Making the Connection: Social Bonding in Courtship Situations. *Amer. J. Sociology* 118, 6 (May 2013), 1596–1649. doi:10.1086/670240

[54] Burt L. Monroe, Michael P. Colaresi, and Kevin M. Quinn. 2008. Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict. *Political Analysis* 16, 4 (2008), 372–403. doi:10.1093/pan/mpn018

[55] Viet-An Nguyen, Jordan Boyd-Graber, Philip Resnik, Deborah A. Cai, Jennifer E. Midberry, and Yuanxin Wang. 2014. Modeling topic control to detect influence in conversations using nonparametric topic models. *Machine Learning* 95, 3 (June 2014), 381–421. doi:10.1007/s10994-013-5417-9

[56] Vlad Niculae and Cristian Danescu-Niculescu-Mizil. 2016. Conversational Markers of Constructive Discussions. In *Proceedings of NAACL.* doi:10.18653/v1/N16-1070

[57] Richard A. Page and Joseph L. Balloun. 1978. The Effect of Voice Volume on the Perception of Personality. *The Journal of Social Psychology* 105, 1 (June 1978), 65–72. doi:10.1080/00224545.1978.9924091 Publisher: Routledge _eprint: https://doi.org/10.1080/00224545.1978.9924091.

[58] Mark T. Palmer. 1989. Controlling conversations: Turns, topics and interpersonal control. *Communication Monographs* 56, 1 (1989), 1–18. doi:10.1080/03637758909390246

[59] Jennifer S Pardo, Adelya Urmanche, Hannah Gash, Jaclyn Wiener, Nicholas Mason, Sherilyn Wilman, Keagan Francis, and Alexa Decker. 2019. The Montclair Map Task: Balance, Efficacy, and Efficiency in Conversational Interaction. *Language and Speech* 62, 2 (June 2019), 378–398. doi:10.1177/0023830918775435 Publisher: SAGE Publications Ltd.

[60] Juan Pablo Pardo-Guerra and Prithviraj Pahwa. 2022. The extended computational case method: A framework for research design. *Sociological methods & research* 51, 4 (2022), 1826–1867. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

[61] Joonsuk Park and Claire Cardie. 2018. A Corpus of eRulemaking User Comments for Measuring Evaluability of Arguments. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018).* European Language Resources Association (ELRA), Miyazaki, Japan. https://aclanthology.org/L18-1257

[62] Vinodkumar Prabhakaran, Ajita John, and Dorée D. Seligmann. 2013. Who Had the Upper Hand? Ranking Participants of Interactions Based on Their Relative Power. In *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, Ruslan Mitkov and Jong C. Park (Eds.). Asian Federation of Natural Language Processing, Nagoya, Japan, 365–373. https://aclanthology.org/I13-1042

[63] Andrew Reece, Gus Cooney, Peter Bull, Christine Chung, Bryn Dawson, Casey Fitzpatrick, Tamara Glazer, Dean Knox, Alex Liebscher, and Sebastian Marin. 2023. The CANDOR corpus: Insights from a large multimodal dataset of naturalistic conversation. *Science Advances* 9, 13 (March 2023), eadf3197. doi:10.1126/sciadv.adf3197

[64] Stuart Reeves. 2017. Some conversational challenges of talking with machines. Portland, Oregon, USA. https://eprints.nottingham.ac.uk/40510/

[65] Felicia Roberts, Alexander L. Francis, and Melanie Morgan. 2006. The interaction of inter-turn silence with prosodic cues in listener perceptions of "trouble" in conversation. *Speech Communication* 48, 9 (Sept. 2006), 1079–1093. doi:10.1016/j.specom.2006.02.001

[66] Harvey Sacks, Emanuel A. Schegloff, and Gail Jefferson. 1974. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 50, 4 (1974), 696–735. doi:10.2307/412243 Publisher: Linguistic Society of America.

[67] Emanuel A Schegloff. 1986. The routine as achievement. *Human studies* 9, 2 (1986), 111–151. Publisher: Springer.

[68] Emanuel A Schegloff. 1993. Reflections on quantification in the study of conversation. *Research on language and social interaction* 26, 1 (1993), 99–128. Publisher: Taylor & Francis.

[69] Emanuel A. Schegloff. 2000. Overlapping Talk and the Organization of Turn-Taking for Conversation. *Language in Society* 29, 1 (2000), 1–63. https://www.jstor.org/stable/4168983 Publisher: Cambridge University Press.

[70] Emanuel A Schegloff. 2007. *Sequence organization in interaction: A primer in conversation analysis I.* Vol. 1. Cambridge university press.

[71] Emanuel A Schegloff, Gail Jefferson, and Harvey Sacks. 1977. The preference for self-correction in the organization of repair in conversation. *Language* 53, 2 (1977), 361–382. Publisher: Linguistic Society of America.

[72] Emanuel A Schegloff and Harvey Sacks. 1973. Opening up closings. (1973). Publisher: Walter de Gruyter, Berlin/New York Berlin, New York.

[73] Emanuel A Schegloff, H Sacks, and G Jefferson. 1992. Lectures on conversation.

[74] Jeremiah Scholl, John McCarthy, and Rikard Harr. 2006. A comparison of chat and audio in media rich environments. In *Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work.* ACM, Banff Alberta Canada, 323–332. doi:10.1145/1180875.1180925

[75] Ronald Scollon and Suzanne Wong-Scollon. 1991. Athabaskan-English Interethnic Communication. In *Cultural Communication and Intercultural Contact.* Routledge. Num Pages: 28.

[76] Joseph Seering, Tianmi Fang, Luca Damasco, Mianhong 'Cherie' Chen, Likang Sun, and Geoff Kaufman. 2019. Designing User Interface Elements to Improve the Quality and Civility of Discourse in Online Commenting Behaviors. In *Proceedings of CHI (CHI '19).* 1–14. doi:10.1145/3290605.3300836

[77] Maya Sen and Omar Wasow. 2016. Race as a Bundle of Sticks: Designs that Estimate Effects of Seemingly Immutable Characteristics. *Annual Review of Political Science* 19, Volume 19, 2016 (May 2016), 499–522. doi:10.1146/annurev-polisci-032015-010015 Publisher: Annual Reviews.

[78] Stephen M. Shapiro. 1984. Oral Argument in the Supreme Court of the United States. *Catholic University Law Review* 33, 3 (1984), 27. 00043.

[79] Jack Sidnell and Tanya Stivers. 2012. *The handbook of conversation analysis.* John Wiley & Sons.

[80] Georg Simmel. 1902. The number of members as determining the sociological form of the group. I. *American journal of Sociology* 8, 1 (1902), 1–46. Publisher: University of Chicago Press.

[81] Bruce L. Smith, Bruce L. Brown, William J. Strong, and Alvin C. Rencher. 1975. Effects of Speech Rate on Personality Perception. *Language and Speech* 18, 2 (April 1975), 145–152. doi:10.1177/002383097501800203 Publisher: SAGE Publications Ltd.

[82] Susan Sprecher, Stanislav Treger, Joshua D. Wondra, Nicole Hilaire, and Kevin Wallpe. 2013. Taking turns: Reciprocal self-disclosure promotes liking in initial interactions. *Journal of Experimental Social Psychology* 49, 5 (Sept. 2013), 860–866. doi:10.1016/j.jesp.2013.03.017

[83] Laura Stafford, Cynthia S. Burggraf, and William F. Sharkey. 1987. Conversational Memory: The Effects of Time, Recall, Mode, and Memory Expectancies on Remembrances of Natural Conversations. *Human Communication Research* 14, 2 (Dec. 1987), 203–229. doi:10.1111/j.1468-2958.1987.tb00127.x

[84] Tanya Stivers. 2021. Is conversation built for two? The partitioning of social interaction. *Research on Language and Social Interaction* 54, 1 (2021), 1–19. Publisher: Taylor & Francis.

[85] T. Stivers, N. J. Enfield, P. Brown, C. Englert, M. Hayashi, T. Heinemann, G. Hoymann, F. Rossano, J. P. de Ruiter, K.-E. Yoon, and S. C. Levinson. 2009. Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences* 106, 26 (June 2009), 10587–10592. doi:10.1073/pnas.0903616106

[86] Elizabeth Stokoe. 2024. Elizabeth Stokoe: "Conversation analysis has many applications, from Silicon Valley to medicine". https://blogs.lse.ac.uk/businessreview/2024/01/29/elizabeth-stokoe-conversation-analysis-has-many-applications-from-silicon-valley-to-medicine/

[87] Deborah Tannen. 1981. Indirectness in discourse: Ethnicity as conversational style*. *Discourse Processes* 4, 3 (July 1981), 221–238. doi:10.1080/01638538109544517 Publisher: Routledge _eprint: https://doi.org/10.1080/01638538109544517.

[88] Deborah Tannen. 2005. *Conversational style : analyzing talk among friends.* Oxford University Press, New York.

[89] Samuel Hardman Taylor, Dominic DiFranzo, Yoon Hyung Choi, Shruti Sannon, and Natalya N. Bazarova. 2019. Accountability and Empathy by Design: Encouraging Bystander Intervention to Cyberbullying on Social Media. In *Proceedings of CSCW*. 118:1–118:26. doi:10.1145/3359220

[90] Emma M Templeton, Luke J Chang, Elizabeth A Reynolds, Marie D Cone LeBeaumont, and Thalia Wheatley. 2022. Fast response times signal social connection in conversation. *Proceedings of the National Academy of Sciences* 119, 4 (2022), e2116915119. Publisher: National Acad Sciences.

[91] Joshua R. Tyler and John C. Tang. 2003. When Can I Expect an Email Response? A Study of Rhythms in Email Usage. In *ECSCW 2003*, Kari Kuutti, Eija Helena Karsten, Geraldine Fitzpatrick, Paul Dourish, and Kjeld Schmidt (Eds.). Springer Netherlands, Dordrecht, 239–258. doi:10.1007/978-94-010-0068-0_13

[92] Fernanda B. Viégas and Judith S. Donath. 1999. Chat circles. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI '99)*. Association for Computing Machinery, New York, NY, USA, 9–16. doi:10.1145/302979.302981

[93] Marilyn Walker and Steve Whittaker. 1990. Mixed Initiative in Dialogue: An Investigation into Discourse Segmentation. In *Proceedings of ACL*.

[94] Deanna Wilkes-Gibbs and Herbert H. Clark. 1992. Coordinating Beliefs in Conversation. *Journal of Memory and Language* 31, 2 (April 1992), 183–194. https://www.proquest.com/docview/1297341531/citation/1718513DD4D748EEPQ/1 Num Pages: 12 Place: New York, United States Publisher: Academic Press.

[95] Michael Yeomans, F Katelynn Boland, Hanne K Collins, Nicole Abi-Esber, and Alison Wood Brooks. 2023. A practical guide to conversation research: How to study what people say to each other. *Advances in Methods and Practices in Psychological Science* 6, 4 (2023), 25152459231183919. Publisher: Sage Publications Sage CA: Los Angeles, CA.

[96] Justine Zhang, Cristian Danescu-Niculescu-Mizil, Christina Sauper, and Sean J. Taylor. 2018. Characterizing Online Public Discussions through Patterns of Participant Interactions. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW (Nov. 2018), 198:1–198:27. doi:10.1145/3274467

[97] Justine Zhang, Sendhil Mullainathan, and Cristian Danescu-Niculescu-Mizil. 2020. Quantifying the Causal Effects of Conversational Tendencies. *Proceedings of the ACM on Human-Computer Interaction* 4, 2 (Oct. 2020), 1–24. doi:10.1145/3415202

[98] Rui Zhang, Nathan J. McNeese, Guo Freeman, and Geoff Musick. 2021. "An Ideal Human": Expectations of AI Teammates in Human-AI Teaming. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW3 (Jan. 2021), 246:1–246:25. doi:10.1145/3432945

[99] Dean H. Zimmerman and Candace West. 1996. 9. Sex roles, interruptions and silences in conversation. In *Towards a Critical Sociolinguistics*. John Benjamins, 211. https://www.jbe-platform.com/content/books/9789027299697-cilt.125.12zim

A Appendix

A.1 Speaker Perception Across Axes

We analyze the variations in enjoyment rating reported by speakers across the blue-red and blue-gray axis through controlling the amount of gray windows to be below 40% and the amount of red windows to be below 5% respectively. The results are presented in Figure 10 and Figure 11. We observe limited difference in speaker perception along the blue-gray axis, and the result for the blue-red axis are similar to that for the gray-red axis shown in Figure 7.
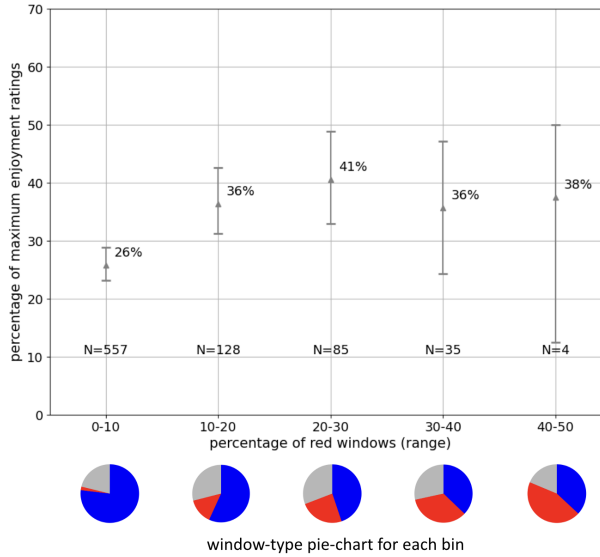

Fig. 10. Similar to Figure 7, with similar percentage of windows that are neutral (gray < 5%), we show difference in speaker perception between conversations with more windows dominated by the secondary speaker (red) and conversations with more windows dominated by the primary speaker (blue).

A.2 Age and Gender

We present the distribution of conversations with varying talk-time sharing dynamics across people from different demographic groups, specifically in terms of age and gender. For each talk-time sharing dynamic stereotype, we show the percentage of conversations that occurs between two males, two females, or a male and a female (Table 5). A similar analysis is conducted for age groups, examining conversations between two younger speakers, two older speakers, and a younger and an older speaker (Table 6). Younger and older speakers are classified based on the median age of all participants, which is 31 years old.

A.3 Measuring speak-time from audio vs. transcripts

We present the validation of transcript accuracy in capturing speak-time by comparing it against extracted timestamps that we extract from the raw audio recordings using a voice activity detection tool, specifically the webrtcvad package in python. We consider the speaking time overlaps between automated transcript and voice activity detection result for every individual conversation speaker. Among all 1594 conversations we used in our analysis (3188 audio files, 2 for each conversation separated by speakers), we observe a median f1 score of 0.86 (0.93 precision and 0.81 recall) across all speaker audio files, with only 6.88% having f1 score less than 0.7. Figure 12 and figure 13 present
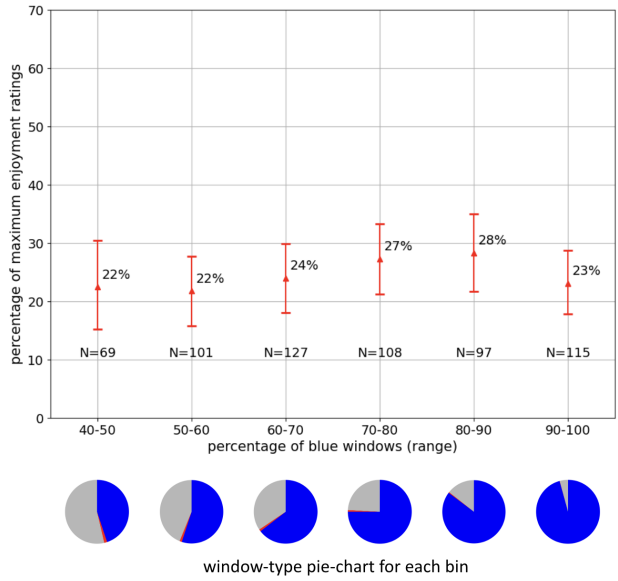
window-type pie-chart for each bin

Fig. 11. Similar to Figure 7, with similar percentage of windows dominated by the secondary speaker (red < 5%), we show difference in speaker perception between conversations with more windows in neutral (gray) and conversations with more windows dominated by the primary speaker (blue).

Table 5. Percentage of conversations with speaker from certain gender group (male, female, or others) across different talk-time sharing dynamic stereotypes.

	male and female	both male	both female	others
overall distribution	48.9%	19.8%	28.0%	3.3%
dominating throughout	49.2%	20.7%	27.1%	3.0%
back and forth	53.4%	19.5%	24.3%	2.8%
alternating dominance	45.8%	18.8%	30.2%	5.2%

two visualizations of the alignment between CANDOR's automated transcript timestamps and the voice activity detection results for the audio recordings of two conversation speakers. We observe a high alignment between the two for most conversations, such as the one shown in Figure 12, particularly in terms of the true positive rate (resulting in a high precision score as mentioned above). However, in rare cases we also notice that additional speech activities detected from the audio signal, as shown in Figure 13. Based on a manual investigation of these cases, these differences can be largely attributed to noise or non-speech sounds, such as laughter or coughing. This results in a lower recall comparing to precision score across audio files. Nevertheless, these "noise" signals could provide additional information to support future analyses. For example, the frequency of laughter could be used to infer participants' experiences during the conversation. Future work could explore additional uses of the audio files to leverage richer signals within.

Table 6. Percentage of conversations with speaker from certain age group (younger than 31 years old or older than 31 years old) across different talk-time sharing dynamic stereotypes.

	younger and older	both younger	both older
overall distribution	51.1%	21.1%	27.9%
dominating throughout	56.3%	18.6%	25.1%
back and forth	50.2%	29.1%	20.7%
alternating dominance	53.1%	17.7%	29.2%
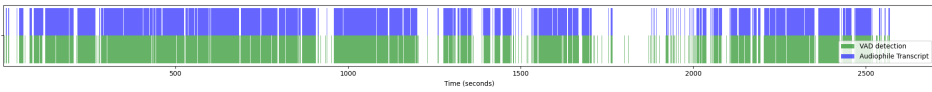
Fig. 12. Visualization on the alignment of timestamps between automated transcript timestamps and voice activity detection results for an individual conversation speaker audio file (CANDOR ID: 68e5b791-3931-440e-a603-4e3a063d8521, speaker ID: 5f880d491c31fc25c442fa70). For most of the CANDOR audio files, we observe good alignment to validate the accuracy of automated transcript timestamps.
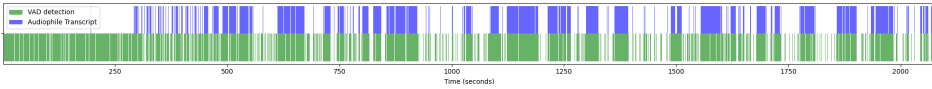

Fig. 13. Visualization of a conversation audio file with bad alignment of timestamps between automated transcript timestamps and voice activity detection results (CANDOR ID: dbbf2d8f-eaba-478c-bf32-330620a4b4ee, speaker ID: 5a70d87e9cdd180001776440). The error in poorly aligned timestamps is hypothesized to be with the problem of voice activity detection's inability to distinguish between speaking language and noise or backchannel voices, which results in extra detected speaking time at the beginning of the conversation that are likely to be noise in speaker's waiting time for this audio.

To validate our results from the exploration of CANDOR corpus, we conduct all analysis again directly on the audio files with voice activity detection algorithm extracted timestamps. We observe same qualitative trend from our main analysis, with slightly more balanced talk-time sharing behavior across conversation as shown in Figure 14. This slight change could be the result of the extra non-language activities discussed above. We take this into account by adjusting the window-dominance parameter accordingly ($M = 55\%$) and present the results in Figure 15, 16, and 17.

A.4 Parameters for Supreme Court Oral Arguments and additional results

Respondent side results. Here we present results from respondent's turn of the oral arguments. As shown in Figure 18, a similar trend is observed. When lawyers dominate throughout oral arguments, they have a greater chance of winning; otherwise, they are more likely to lose the case. The difference is more pronounced in unanimous cases, in which the respondents are substantially less likely to win.
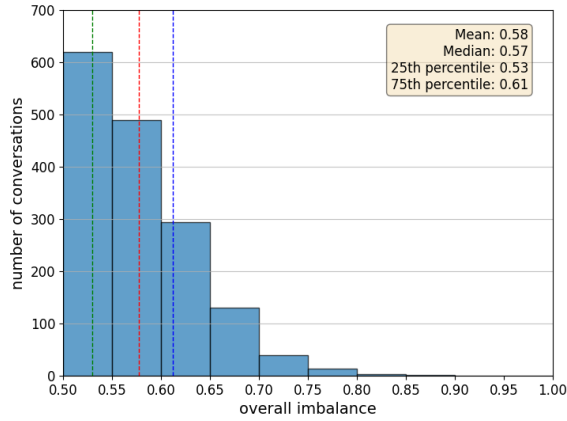
Fig. 14. Distribution of conversation-level imbalance in the CANDOR corpus with timestamps directly extracted from raw audio files using voice activity detection tool.
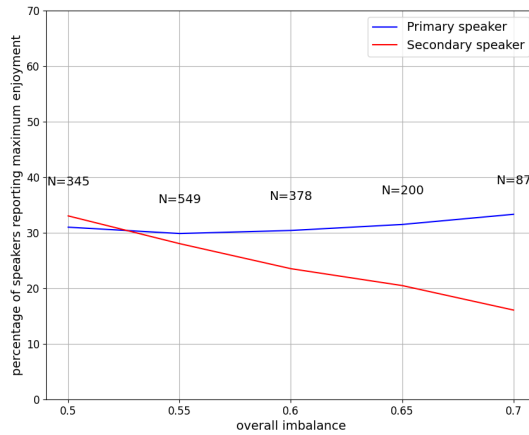

Fig. 15. Similar to Figure 3 but with speaking timestamps captured using voice activity detection tool, we show the percentage of primary and secondary speakers reporting maximum enjoyment scores for conversations with different levels of imbalance. As in Figure 3, levels with 30 or less conversations ($N < 30$) are not shown. The same qualitative trend is observed.

A.5 Additional Examples

In this section, we present additional visualizations and examples of speaker comments to further support our qualitative interpretation.

More conversation visualization showcases. We present more conversation talk-time dynamics visualization from conversations in the CANDOR corpus. Figure 19 shows the visualization of talk-time dynamics for 54 conversations sorted by their conversation-level imbalance. Note that while
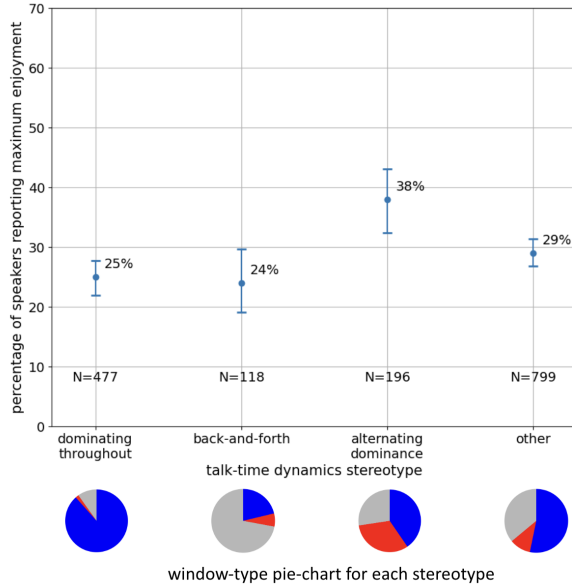
Fig. 16. Similar to Figure 6 but with speaking timestamps captured using voice activity detection tool, we compare the reported enjoyment of the three stereotypical talk-time dynamics: dominating throughout, back-and-forth, and alternating dominance. The same qualitative trend is observed.
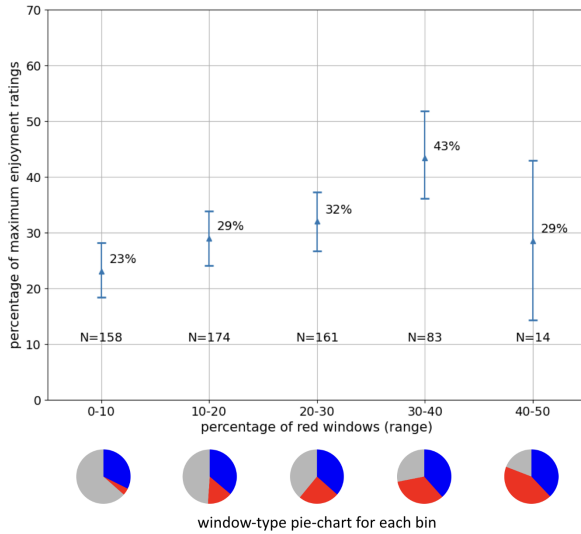

Fig. 17. Similar to Figure 7 but with speaking timestamps captured using voice activity detection tool, we control for the proportion of blue windows, and compare reported enjoyment for varying proportions of red windows. The same qualitative trend is observed.

more imbalance conversations are largely in dominating throughout stereotype, more balanced conversations can be either back-and-forth or alternating dominance.
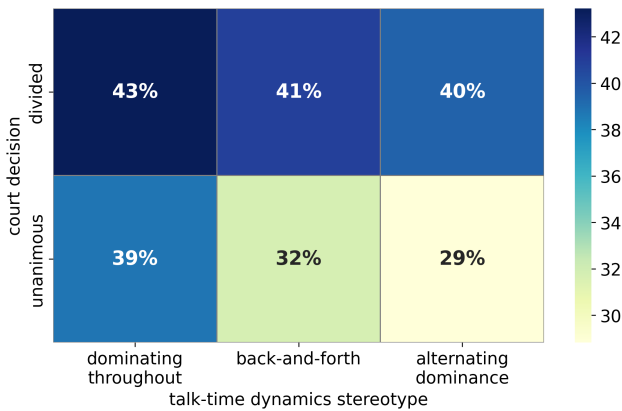
Fig. 18. Comparing the respondent' winning rate across different stereotypes of conversations and in different court decision types.

Combined stereotype conversations visualization showcases. We present visualization from conversations with a combination of stereotypes as described in Section 4. Figures 20, 21, 22, 23, and 24 shows all 46 conversations with a combination of stereotypes captured by our framework.

A.5.1 More User Comment Examples. In this section, we provide expanded versions of Table 2, 3, and 4 including the full speaker comments to give a more complete picture of how users describe their experiences. See Table 7, 8, 9, 10, 11, and 12.

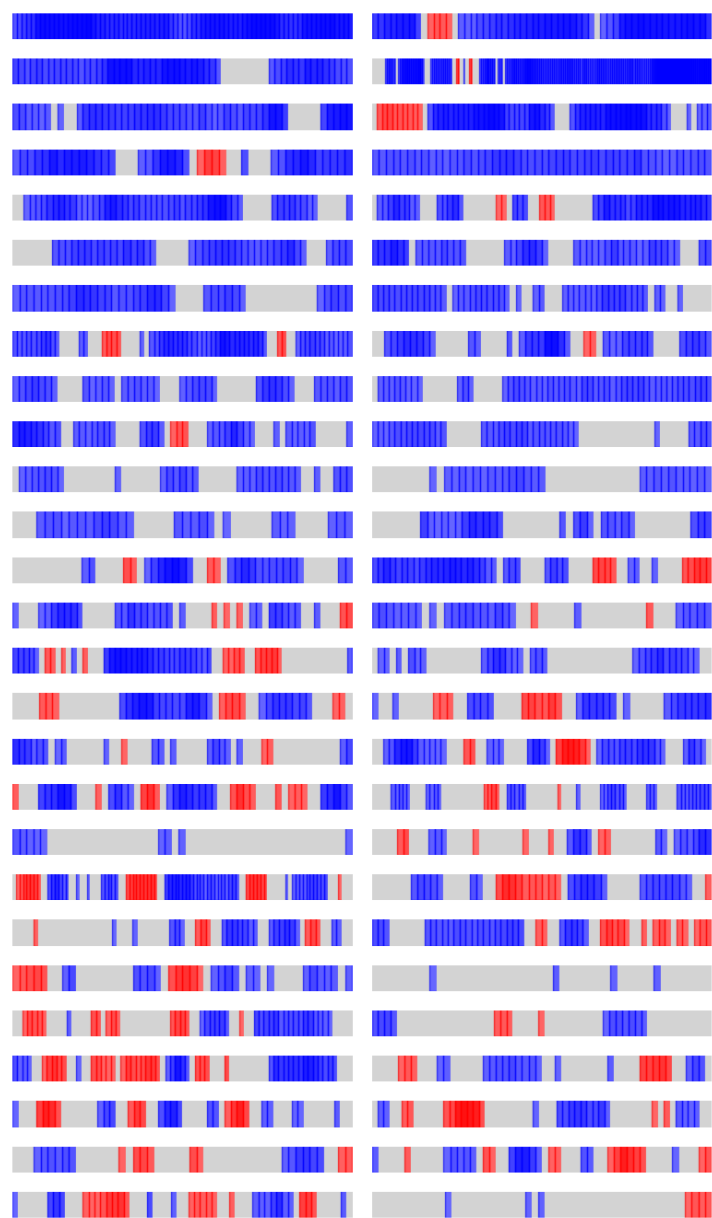Received July 2024; revised December 2024; accepted March 2025

Fig. 19. Visualization of talk-time dynamics of 54 conversations in the CANDOR dataset, sorted in descending conversation level imbalance.
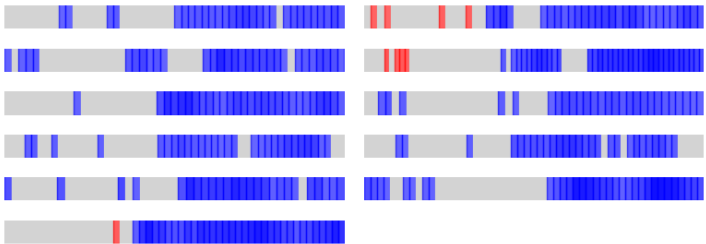
Fig. 20. Back-and-forth transitioning to dominating throughout conversations
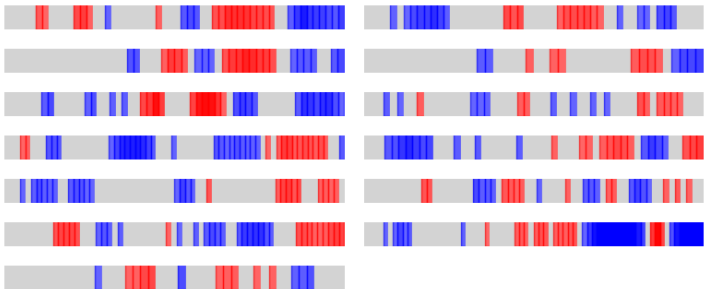

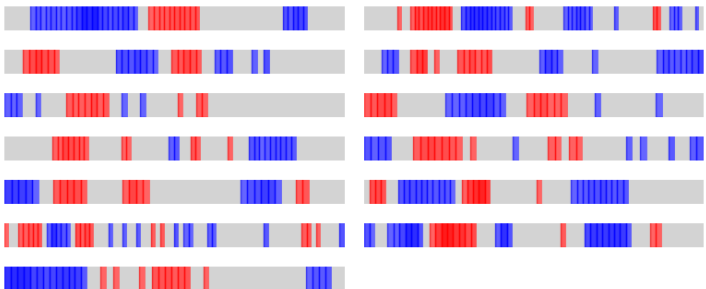Fig. 21. Back-and-forth transitioning to alternating dominance conversations

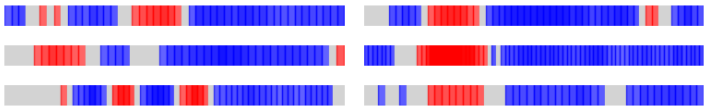Fig. 22. Alternating dominance transitioning to back-and-forth conversations

Fig. 23. Alternating dominance transitioning to dominating throughout conversations
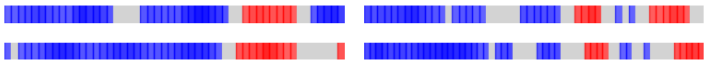
Fig. 24. Dominating throughout transitioning to alternating dominance conversations

Table 7. Extended version of Table 2: phrases from positive comments that most distinguish between balanced and imbalanced conversations, with representative excerpts.

	top distinctive phrases	example excerpts
Positive comments for **highly-balanced** conversations	both, lot in common, were able, were able to, like we, lot in, about our, able, able to, we both	• **We both** had a **lot in common**, so it made talking with each other super easy, and almost **like we** knew each other. • We had a lot to talk about, and seemed to have a **lot in common**. **We both** had single children and **were able to** connect with that. • We **were able to** connect well over our experiences with prolific and the study as well as sharing similar hobbies and likes. • I felt **like we** listened to one another. • I feel **like we** connected and that makes it easier. • It was interesting learning **about our** similarities and differences in ideas. • I felt open and **able to** share what is going on in my city and life.
Positive comments for **highly-imbalanced** conversations	chat, he, partner, talked lot, listener, else, im, is, she, bring	• I feel like my **chat partner** was great at asking me questions and keeping me talking. • It seems like **he** is a talker in real life and **he** likes to be the center of attention, so maybe this helped in our video **chat** that we had. • **She talked a lot** and was very pleasent and smiled a lot. • Although I tend to be an introvert, I know a **talked a lot** and had to intentionally stop myself and ask him questions and invite him to talk more. • **She** was also relaxed, an active **listener**, and asked good questions. • Fortunately, I am a good **listener**. • I prefer when the other person **brings** up conversation topics because it is sometimes difficult for me to come up with things to say.

Table 8. Extended version of Table 2: phrases from negative comments that most distinguish between balanced and imbalanced conversations, with representative excerpts.

	top distinctive phrases	example excerpts
Negative comments for **highly-balanced** conversations	our, **our conversation**, think of, of anything that, we had, anything that, both, her and, think the conversation, each other	• Honestly, **our conversation** was top notched! We hit it off right away and had no issues reaching the 25 minute mark. • We flowed easily in **our conversation** and I am typing to get to fifty word count. • I think the rough parts were mainly how we tried to **think of** how to go back and forth, at first. • I can't really **think of** anything that went wrong. • We **both** tend to be more introverted as we mentioned in the conversation. • I felt like I was under pressure to just agree with **her and** nod along even when I was not interested. • We spent twice the recommended time chatting and I could totally see being friends with her if we were closer to **each other**.
Negative comments for **highly-imbalanced** conversations	than, **more than**, much, talked, me, too much, talked too, talked too much, he, ended up, ended	• I think **he** did talk a fair amount **more than** me, which was mildly frustrating, but **he** meant well and I understood it better when **he** told me **he** was originally from New York. • **He talked** a lot **more than** I did. • I guess one thing was that she did most of the talking, but I was happy to listen to someone who knows **much more than** I do. • I think I **talked too much**. • I probably **talked too much**, but I felt like she wasn't giving very much in return so I don't think I had a choice. • I think she felt that way too and we **ended up** sort of slowing down and stretching out to cover the time needed. • **He ended up** opening up toward the end and actually provided me with some information that I might find useful for everyday life as well.

Table 9. Extended version of Table 3: phrases that most distinguish between positive comments by primary and secondary speakers for highly-imbalanced conversations.

	top distinctive phrases	example excerpts
Positive comments from **primary speaker**	questions, asked, but, probably, us, about my, agreed, some, young, good listener, each, interested	•She was a **good listener** and **asked** pertinent **questions** to get me talking. • My high energy and inquisitive nature allowed me to ask many **questions** of the other participant. • The way he was curious and listened and **asked** more and more **questions**. • I think that both of **us** felt awkward at first, but she did a great job of normalizing that and she seemed to have a lot to talk about. • I probably did the majority of the talking, but they **asked** good **questions** and gave interesting answers when I **asked** them **questions**. • I was able to learn a lot **about my** partner without being prying on invasive. • She was also fun to talk to because she's a **good listener** and was understanding toward my situation.
Positive comments from **secondary speaker**	had lot, something, lot, stories, had similar, had lot of, share, friendly and, that my, life	•He **had a lot** to talk about so I let him tell me about those things for as long as he wanted to. • We **had a lot** in common even though we were very different people. • We found **something** that we could connect on right away. • He had great life experience and **stories** that he **shared**. • We also **had similar** experiences in some ways, which helped out. • My partner was very chatty and willing to **share** her views, while I was quite willing to listen. • We both were **friendly and** eager to listen to each other.

Table 10. Extended version of Table 3: phrases that most distinguish between negative comments by primary and secondary speakers for highly-imbalanced conversations.

	top distinctive phrases	example excerpts
Negative comments from **primary speaker**	her, little, too much, we had, flow, talked too, should, she was, she is, shy	•My partner was not the best at **her** side of the discussion, sometimes failing to move the conversation forward and creating awkward silences. • I felt a **little** scattered about my answers. • I think **she was** a **little** quiet by nature which is ok. • The conversation felt somewhat forced, **we had** to find things to talk about. • I think I probably **talked too much**. • I felt I led the conversation **too much** and perhaps dominated. • He was kind of **shy**, and quiet. • She seemed very distracted by something else, which made it difficult for the conversation to **flow**.
Negative comments from **secondary speaker**	different, hard, say, went, the conversation was, talked lot, life, honestly, didnt, kept, it was hard	•I think we are just two very **different** people, so we **didn't** have a lot in common. • He talked way too much; **it was hard** to get a word in edgewise. • I did not know what to **say**. • What **didn't** work in **the conversation was** the lack of equity in conversation talking time. • My partner **talked a lot** about herself and **didn't** ask many questions, which made it harder for me to share as well. • He **talked a lot** more than I did. • In real **life** I tend to be on the quiet side so this was expected.

Table 11. Extended version of Table 4: phrases that most distinguish between positive comments for stereo-typical alternating dominance and back-and-forth conversations.

	top distinctive phrases	example excerpts
Positive comments for **alternating dominance** conversations	stories, to listen, listen, stranger, life experiences, the other, myself, his, good conversation, super	• We both had interesting **stories** to tell, so that was also very enjoyable. • She had some really interesting **life experiences** and **stories** to share. • I feel like we were both able **to listen** actively to one another, and bring curiosity to the conversation. • We both listened when **the other** was speaking and asked follow-up questions. • I felt that that the conversation did not feel like I was talking to a complete **stranger**. • I felt comfortable talking to **the other** person and was willing to open up about **myself**, which I think led to a **good conversation** overall. • We had a **good conversation** about several topics about movies, books, and work.
Positive comments for **back-and-forth** conversations	which, the conversation going, conversation going, topics, keep the, keep the conversation, games, going, well was, discussed	• We both took turns talking and **keeping the conversation going** with each other. • He **kept the conversation going** at a good pace, and really was able to steer the conversation back when I ventured off **topic**, or was obviously uncomfortable. • The flow of the conversation went well in which we kept bringing up **topics** to speak on and it went smoothly. • I think the wide variety of **topics** we **discussed** was interesting and informative. • Not a lot outside of talking about **games**. • So we **discussed** many of our favorite **games** together and talked about our emotional responses to the **games**. • What worked **well was** that my partner was willing to keep talking.

Table 12. Extended version of Table 4: phrases that most distinguish between negative comments for stereo-typical alternating dominance and back-and-forth conversations.

	top distinctive phrases	example excerpts
Negative comments for **alternating dominance** conversations	go, one, interested, she, into the, minutes, get, wasnt really, through, few minutes	• I don't **really** have much to say about what didn't **go** well. • My partner didn't seem **interested** in me at all. • I hope that **she** felt able, and had the space to direct the conversation along lines that **interested** her. • He did a good job listening to me when I talked, but sometimes he didn't easily let me **into the** conversation. • As soon as we were a **few minutes** over, **she** wanted to end the conversation. • I fell guarded the first **few minutes** of the conversation because I did not know how to proceed with this person. • **She** laughed at stuff that **wasn't really** funny several times, so I couldn't tell if it was for humor or nerves.
Negative comments for **back-and-forth** conversations	some, going, im, most, had to, went well, understand, flow, kept, topics, moments	• The conversation had **some** awkward **moments** and silent pauses. • Since we were both a bit more shy, we had **some moments** of dead air. • There were **some moments** where I wasn't sure what to talk about next. • **Most** of the conversation did work well, but I found I **had to** do most of the prompting at the start of the conversation. • He didn't have a lot of talking points and I felt as though I **had to** keep asking questions to keep the conversation **going**. • The conversation **flow** seemed forced especially when one topic had been discussed and we need to find another. • The communication, I couldn't always **understand** him, as I said his accent was pretty strong for me to follow and always **understand**. • There were times where it felt like I was interrupting what he was trying to say or I didn't fully grasp or **understand** what he was saying.