Profiling Turns in Interaction: Discourse Structure and Function

Sherri L. Condon
Department of English
Center for Advanced Computer Studies
Institute of Cognitive Science
University of Louisiana at Lafayette
condo@louisiana.edu

Claude G. Čech
Department of Psychology
Center for Advanced Computer Studies
Institute of Cognitive Science
University of Louisiana at Lafayette
cech@louisiana.edu

Abstract

Turn-taking provides a basis for comparing interactions in different communication environments, and this paper demonstrates that readily observable features of turns can be linked to principles that organize and manage the interaction. Results are based on 150 decision-making interactions elicited in a face-to-face environment, an asynchronous, e-mail environment, and several types of synchronous computer-mediated environments. We show that three features of turns can be linked to discourse structure and function. First, as turns increase in size, participants switch from serial to parallel strategies to organize their decision-making. Second, pivot turns, which are turns that are much shorter than the turns that precede or follow them, can reflect the discourse functions of the relevant turns. Finally, turns can be used for measures of dominance based on turn size. We conclude that designers of communication systems can take advantage of peoples' ability to develop effective strategies for packaging messages in different environments.

1. Introduction

Ordinary face-to-face interaction is designed to satisfy the exigencies of verbal communication by exploiting the sequential dependencies or contingencies that arise as the talk proceeds [15,17]. It is these links between participants' contributions that make conversations truly interactive, and the practice of turn-taking provides a framework for managing interaction that has become central to understanding the organization of language behavior in conversation [12, 14]. Developments in communication technology such as the Unix Talk function make it theoretically possible for participants to interact without discernable turns, but turn structures appear to be important in this medium, too: van der Wege & Clark, studying 22 conversations in the Talk environment, report that less than 3% of the words overlap [19]. Examining turn-taking in different communication environments should facilitate understanding how readily observable, superficial features of turns can be linked to principles that organize and manage the interaction.

In our research program, we ask participants to cooperatively plan events in different communication environments in order to examine the strategies that emerge when they manage the same work and package the same kinds of messages in different media. Computer-mediated communication systems allow researchers to systematically vary features of the communication environment that cannot be altered in face-to-face interaction. For example, face-to-face interaction is always *synchronous*, which we define as communication in which messages are produced with the expectation that they will be processed and answered immediately. Consequently, participants in face-to-face interaction interpret their partners' contributions while simultaneously planning their own next contributions. Moreover, another fixed feature of face-to-face interaction, the fact that simultaneous production of messages results in degradation of the signal, requires that speakers also identify an appropriate location in the flow of talk to produce their contributions. Finally, the lack of permanence or persistence of the message in face-to-face conversation places additional loads on working memory. Precisely-timed exchanges of relatively short messages satisfy the constraints imposed by these processing demands, which may explain the emergence of turn-taking practices [11, 15, p. 75].

Unlike face-to-face interaction, computer-mediated interaction can also be *asynchronous*, as in e-mail. Furthermore, synchronous computer-mediated environments can include systems in which participants' messages appear in a single graphic window on their screens, as in typical chat environments, or systems in which messages appear in separate graphic windows, as in the Unix Talk and Ytalk environments. Systems may also be devised in which the screen displays just a single message, simulating some of the signal degradation that occurs with simultaneous production in verbal interactions, as in Condon & Čech [6]. The Fugue system [18] represents temporal relations among participants' contributions like a musical score in the spatial arrangement of messages, and a permanent record of the interaction is accessible through a graphic representation.

Our research suggests that participants in synchronous communication manage the contingency of interaction in ways that might not be possible in asynchronous communication—and vice versa [7]. Similarly, demands on processing will be altered by variation in other features of the communication environment such as message persistence. Processing demands are reflected in the strategies participants adopt to manage the decision tasks [6]. Focusing on turn structure and function provides a framework for comparing these strategies that foregrounds the encoding and packaging work required for successful interaction. Accordingly, we introduce the notion of a turn profile, a 2-dimensional representation of the length and sequence of turns produced by each participant.

We will present several features of turn profiles that can readily be linked to discourse structure and function. Turn size and changes in turn size are easily read off of the y-axis, and we find that turn size reflects fundamental organizing principles of the interaction. For example, as turns lengthen, participants switch from a serial strategy in which few decisions are addressed per turn to a parallel organization in which many decisions are addressed. Observing changes in turn size, we identify turns that are much shorter than the surrounding turns. We define such turns as *pivot turns*, and we explore how and whether these reflect specific discourse functions. Finally, turn profiles can be used to observe the relative contributions of each participant, enabling measures of relative dominance that may be related to features of the communicative interface. We also describe a way to make turn profiles even more informative and propose some consequences of our work for the design of communication systems.

Our research has led us to conclude that the most efficient and effective communication practices will differ across communication systems to reflect differing processing demands. Moreover, interlocutors in these various environments are reasonably adept at discovering productive practices. Awareness of this versatility can encourage designers to develop communication systems that exceed expectations based on the limits of face-to-face processing.

2. Data Collection and Analysis

This report draws on 150 interactions collected in several studies, some of which are still ongoing. In the first study, 32 dyads engaged in simple decision-making by jointly planning a social event (a picnic or barbecue) and an itinerary (such as a weekend trip). Half of these dyads interacted face-to-face, and the other half engaged in synchronous computer-mediated communication. In the second study, we collected 68 dyadic interactions (including 8 face-to-face interactions) involving a more complex decision-making task (planning the MTV Music Video Awards Show). For the 60 computer-mediated interactions in this study, we varied the interface to enable different amounts of viewable incoming or outgoing material. These interactions were evenly divided among conditions in which participants could view 4, 10, or 18 lines (compared to the 3 or so lines of viewable material in the computer-mediated condition of the initial study).

The first corpus was generated by mixed-sex pairs whereas all others were generated by both mixed- and same-sex pairs. A third corpus includes another 20 interactions using the MTV task. In this corpus, including 10 interactions from Gray's study involving groups of three students rather than dyads [10], participants communicate asynchronously via e-mail. Finally, Babineaux's corpus of 20 interactions was elicited by asking participants to complete the MTV task in a synchronous computer-mediated environment that is very different than the ones employed in the first two studies [1]. Excluding the paid participants in Gray's study, all participants were native English speakers at the University of Louisiana who received extra credit in Introductory Psychology classes for their participation.

The dyads who interacted face-to-face sat at a table with a tape recorder, and the pairs in synchronous computer-mediated conditions were seated at microcomputers in separate rooms. In the first two studies, the latter communicated by typing messages which appeared on the sender's monitor as they were typed, but did not appear on the receiver's monitor until the sender pressed the <Enter> key. The software incorporated this feature to provide well-defined turns and to make it possible to capture and change messages in future studies. In addition, to minimize message permanence and more closely approximate face-to-face interaction, the screen displays text from only one message at a time: messages disappear as soon as the partner begins to type a response. In contrast, the synchronous system employed in Babineaux's study includes a separate, 10-line message area for each participant in which messages appear in real time, as they are typed. This system thus resembles the Unix Talk environment.

The few differences in the computer-mediated conditions of the first two studies include differences in the arrangement of information on the screen, such as a brief description of the MTV problem at the bottom of the screen and the use of an answer form in the first study, but not the second. For more details about the communication interfaces in the two studies see [2,6] Participants in the synchronous interactions were allotted two hours to complete the tasks, and most finished within two hours, though a few computer-mediated interactions were longer. Participants in the e-mail conditions, who were permitted to use any e-mail system they preferred, were asked around midterm to finish by the end of the semester. They were also asked to send us a copy of every message they contributed.

Face-to-face interactions were transcribed from audio recordings into computer files (see [5] for coding and transcription conventions). All interactions were divided into utterance units, defined as single clauses with all complements and adjuncts, including sentential complements and subordinate clauses. Interjections and discourse markers like *yeah, now, well,* and *ok* are treated as separate utterances. Utterance units are annotated to reflect their discourse functions according to a scheme documented in the training manual for coders. Coders are regularly tested using standard annotations prepared by the first author throughout training and subsequently using their coded transcripts. All transcripts were coded by at least one coder who regularly scored above 90% agreement with the standard for all categories, except those in Gray's and Babineaux's corpora, for which these measures are unavailable. Reflecting the multifunctionality of language, our coding scheme assigns every utterance a code from each of 5 groups. Four of these are presented in Table 1; a 5th group provides the "Closing" annotation for utterances that formulate closing bids or function as closings and annotations for spatial features of computer-mediated texts

In addition to analyzing discourse functions, we also analyze turn and utterance size**,** although these measures are complicated by several factors. One factor concerns how to identify words in turns. For example, software that recognizes word boundaries just using spaces and punctuation will count *R&B* ("Rhythm and Blues") as one word, whereas software sensitive to the ampersand will count *R&B* as three words. The preferred approach is not obvious. Clearly, decisions about word recognition will affect the results presented, though the effects will generally be small. Moreover, identifying turn boundaries is not always easy: speech occasionally overlaps in face-to-face interactions, as do messages in the computer-mediated corpus that allows simultaneous transmission (Babineaux's). Ironically, the importance of turn-taking in conversation does not ensure that interaction consists entirely of well-defined turns. Instead, careful analysis of overlapping speech in conversation reveals systematic uses for overlaps such as showing enthusiasm [16].

Faced with problematic turn boundaries, we stress that our transcription and annotation systems are measuring devices and focus on making our measures consistent, establishing conventions that are arbitrary in order to place text in neat sequences of annotated utterances. How these decisions relate to fundamental claims about language use are empirical questions that require further research. Of course, we also try to use any available information to determine utterance and turn boundaries, including prosodic features such as pauses and syntactic features such as completion of phrases and sentences.

In the synchronous computer-mediated environments that do not permit overlap, the design of the communication system provides well-defined turns, but there is an new complication caused by the fact that participants can send more than one message before the partner replies. Consequently, turn boundaries will differ if defined at speaker changes rather than at message boundaries. Below, we employ speaker change as an indicator of turn boundary. Because most message units are also turn units, results are not greatly affected by the choice of unit. For example, the average message size (16.8 words) is less than 2 words smaller than the average turn size (18.6 words) in the 4-line condition. When data is presented, the terms *message* and *turn* will be used to discriminate transmission units and

Table 1: Major Groups of Functions Identified in the Coding System

Move Functions	Response Functions	Other Functions	Explicit Management
Greeting/Salutation	Agrees with Suggestion	Discourse Marker	Task Management
Suggests Action	Disagrees with Suggestion	Orients Suggestion	Verifies Decision
Requests Action	Complies with Request	Source Negotiation	Decision Management
Requests Validation	Acknowledges Only	Closing Formula	Transmission Management
Requests Information	No Clear Response function	Signature	Repair
Articulates a Response		Personal Information	Other Explicit Functions
Elaborates, Repeats		Jokes, Exaggerates	No Clear Explicit Function
No Clear Move function		Name	
		No Clear Other function	

change-of-speaker units respectively.

3. Turn Profiles and Discourse Management

Table 2 presents properties of 120 interactions comprising over 130,000 words. In the e-mail environment, we treat each message as a turn. Two rows are bolded to facilitate comparison of 2 conditions that are of particular interest because dyads completed the same MTV task in reasonably similar total numbers of words, but packaged those words using extremely different strategies. In the first of these, participants interacting face-to-face produced an average 229 turns of less than 10 words per turn, whereas participants in the second, the dyadic e-mail condition**,** averaged 15 turns that were over 100 words in length. Average turn sizes and total turns per interaction in the synchronous computer-mediated conditions clearly lie between the face-to-face and e-mail conditions. Yet the average number of turns in the synchronous computer-mediated interactions is always closer to the average for the e-mail interactions, while the average turn size is closer to the average for the face-to-face interactions. As a result, the MTV task is completed in about half the number of words in the synchronous computer-mediated conditions compared to the face-to-face and e-mail conditions.

Figure 1 presents turn profiles of a face-to-face interaction and an e-mail interaction. In turn profiles, the X-axis represents the order of turns, the Y-axis represents the length in words of each turn, and each participant's turn is represented by a different pattern or color. The extremely different turn packaging strategies are striking in these profiles. Figure 2 presents two profiles from the synchronous computer-mediated MTV conditions that illustrate a preference for short and long turns, respectively. Note the similarity of the "short turn" profile to the face-to-face profile, and of the "long turn" profile to the e-mail profile.

The short vs. long turn packaging corresponds to important differences in the organization of the interactions. In the face-to-face interactions and many of the computer-mediated interactions, each decision is addressed in sequence with an utterance that orients a suggestion by formulating a goal for the decision. The face-to-face excerpt in (1) provides an example from the MTV corpus.

(1) a. P1: who should win best Alternative video.*[Orients]*
 b. P2: Pres. of the united states *[Suggests]*
 c. P1: ok *[Agrees]*
 d. P2: who else should we nominate *[Orients]*
 bush goo-goodolls and oasis *[Suggests]*
 e. P1: sounds good, (...) *[Agrees]*

The routine calls for an utterance with an orienting function like (1a,d) to be followed by an utterance with a suggesting function, as in (1b,d). The suggestion provides a proposal that satisfies the goal formulated in the orientation. The next utterance then agrees with the suggestion, as in (1c,e). (2) provides an excerpt from a synchronous computer-mediated interaction using the same routine.

Table 2: Average Turn Properties in Face to Face (ftf), Synchronous Computer-Mediated (cmc), and E-Mail Environments for Simple (1st) and Complex (MTV) Decision Tasks

Interaction Conditions	Number in Corpus	Average Total Words per Interaction	Average Number of Turns per Interaction	Average Words per Turn per Interaction*
1st ftf	16	1125.6	125.9	6.4
1st cmc	16	311.2	28.7	11.7
MTV ftf	**8**	**1837.3**	**228.7**	**8.6**
MTV 4-line cmc	20	845.1	45.0	18.7
MTV 10-line cmc	20	942.7	48.8	29.2
MTV 18-line cmc	20	794.4	35.1	29.9
MTV e-mail 2 people	**10**	**2266.3**	**14.6**	**149.3**
MTV e-mail 3 people	10	**1833.6**	24.0	77.4

*These figures are obtained by calculating average turn size for each interaction, then averaging those averages

(2) a. P1: who's going to win? *[Orients]* Mariah?*[Suggests]*
 b. P2: yeah probably *[Agrees]*

c. P1: alright Mariah wins what song? *[Orients]*
d. P2: uh Fantasy or whatever? *[Suggests]*
e. P1: that's it that's the same song I was thinking of *[Agrees*] alright alternative *[Orients]* Alanis? *[Suggests]*

As (1) and (2) illustrate, the work of establishing the next decision is accomplished in the orienting utterance. Then the suggestion and agreement routinely follow.

The serial short-turn strategies illustrated in (1) and (2) characterize all of the interactions in the first study. The face-to-face interactions and many of the synchronous computer-mediated interactions in the MTV study also use short-turn strategies. In fact, the five computer-mediated interactions with the shortest average turn sizes averaged 8.4 words per turn, nearly identical to the comparable face-to-face condition. (Of course, the average 67 turns for these five interactions is less than 1/3 the average 229 turns in the face-to-face condition.) Though the computer-mediated interactions contain larger proportions of utterances participating in routine sequences [8] and larger proportions of utterances that explicitly manage the interaction by referring to the interaction itself [6], computer mediation did not change the packaging and decision-making strategies in these interactions.

But computer mediation did have a clear effect in other MTV interactions. The sequence in (3) demonstrates the long turn strategies that emerged.

(3) P1: I agree with you. I'm not sure if you got the message that I just tried to send, but I accidentally pressed "enter" before I had finished. Anyway, so we've got our host. Now we've got to get nominees for our categories. With alternative we could use Bush, Goo Goo Dolls, and Presidents of the United States of America. With pop we could use (but of course) Hootie and the Blowfish, Mariah Carey, and Natalie Merchant. You can pick the nominees for best video. For opening act I say we should get Hootie just because everyone loves them.

P1: (cont.) How many more categories are there?

P2: O.k. on all of that! Best video nominees could possibly be Alanis Morisette's Ironic, Cranberries' Salvation, I am other videos. Can you find two more? There is the categorie of rap, newcomer, best male, and those who will present the nominees and awards. Don't forget there has to be a place to hold the ceremony!

Long turn strategies employ a parallel arrangement of decisions in which orientations, suggestions and agreements for several decisions are encoded in a single turn. P2's first utterance also shows how parallel strategies conserve turns by allowing a single agreement to sanction several suggestions. Also, utterances that encode the routine sequences of orientation, suggestion, and agreement are not adjacent in parallel strategies. Instead, they are usually only contained within adjacent turns.

Short and long turn strategies also exhibit differences in explicit management. (3) includes explicit transmission management near the beginning of P1's contribution, explicit verification of a decision in *Anyway, so we've got our host,* and explicit decision management in *Now we've got to get nominees for our categories* and *You can pick the nominees for best video*. In the interactions with serial short turn strategies, utterances encoding explicit management are rarely combined in the same turn with utterances that encode routine decision-making functions, but with a long turn strategy, all the work is managed in parallel.

The emergence of the parallel long turn affects the cognitive management of the task and loosens the adjacency of utterances linked in routines. Moreover, an interaction like (3) would likely be untenable in face-to-face interactions due to the processing and memory demands described in the introduction. Long turn strategies exploit the fact that long messages can remain on the screen while the reader studies them and plans a response. Though they are synchronous according to our definition, our synchronous computer-mediated environments can be made more asynchronous by taking advantage of this message perma

Figure 1: Turn profiles of a face-to-face interaction (left) and an e-mail interaction (right)

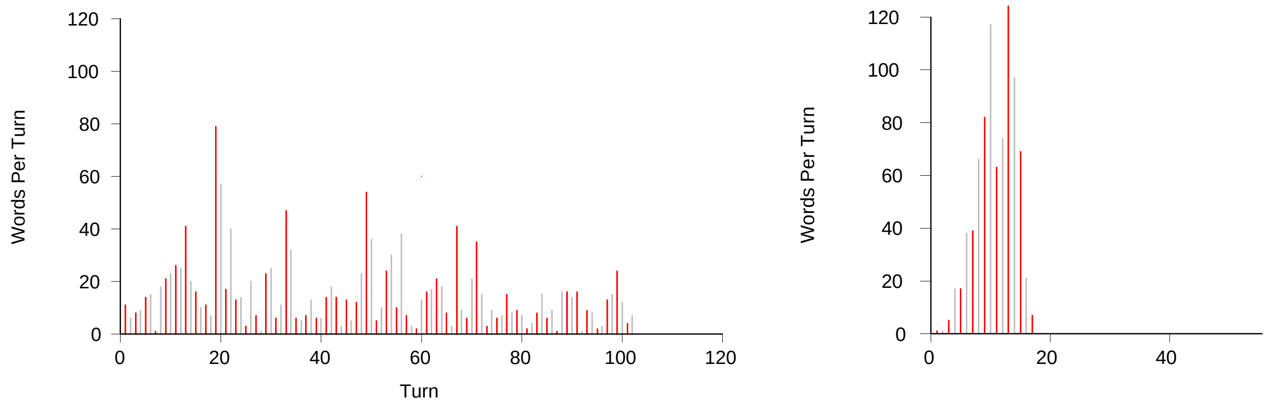


Figure 2: Turn profiles for 2 synchronous computer-mediated interactions illustrating short turn (left) and long turn (right) strategies

nence, which is required for asynchronicity. Some additional evidence that long turns reflect exploitation of the minimal asynchronicity and message permanence available in the system comes from the data Babineaux collected in the fully synchronous Talk-like environment: long turn organization is absent from the corpus.

In this context, it is significant that the parallel long turn strategies resemble the fully asynchronous e-mail interactions in all of the features just described. In the latter, messages are even longer, decisions are also achieved in parallel, and decision-making functions are combined with other management functions, as the short message in (4) illustrates.

(4) Hello! Sorry it took so long for me to respond to your letter..I haven't been at school lately. Vacation was great..but it never seems to last long enough! I think for best female vocalist we should chooose Alanis..seems like she has gotten really hot lately. I think your idea of going from easy listening to heavy is great. I don't have my paper with me right now that has all of our choices on it so there isn't much more I can think of..so I will close here and check in probably on wednesday.. talk to you later

Alicia

A more structured message is illustrated in (5).

(5) HI Traci,

You are never gonna guess where i went Friday. I went to the Bush concert. I felt sort of out of place because i often only listen to country.
My choices for the categories are:
Heavy Metal -- Bush
Alternative -- Alanis
R&B --- Mariah Carey
Is Coolio and Snoop Doggy Dog , one band? If it is, i think we may need one more because we have three for the others. How many bands do you think we need to play during the show? I thought about Alanis playing for one of the bands during the show. What do you think? Maybe also someone that is not a nominee? We also need presenters. I 'm not sure but does Whoppi Goldburg ever present awards for Mtv. I also thought about Ali Landry presenting an award(Miss USA) because she presents awards for other events. It's just a thought, tell me what you think. I really have no idea about High Entergy though. I want to let you know that i'm going back home to my parents for Easter, so i will not be able to reply until next week.While i'm on break i'll be thinking about ideas. Well I hope you have a nice Easter and I'll talk to you next time. Megan

The messages in (4) and (5) illustrate the homogenous turn structure observed in the e-mail data, in which each message is treated as an event to be opened and closed. The opening and closing work, transmission management, interpersonal management and other managerial functions cluster at the beginnings and ends of the messages, while decision-making work occupies the middle. This pattern resembles the one reported by Herring for contributions to an academic listserve [9]. The messages in her corpus are also frequently framed by epistolary devices such as the greetings and signatures in (4) and (5). They begin with introductions that often refer to previous messages or comments using the names of the authors and typically include a closing portion of interpersonal work such as apologies, appeals, and admonitions. In the middle, the "contentful message" of the "body" [9, p. 84] performs the primary or foregrounded work of the contribution, corresponding to the decision-making in our data.

Though asynchronous communication relieves many of the processing demands on fact-to-face interaction, the loss of temporal contingency introduces other pressures. The practices that frame e-mail messages seem to be adaptations to the loss of temporal coincidence and the strategies that depend on it. Participants can no longer rely on swift verification that messages have been received and understood, and delays in transmission increase the likelihood that contextual frames established when the message was formulated are no longer active. Many differences between long turns in the slightly asynchronous computer-mediated conditions and the longer messages in the fully asynchronous conditions seem to be the result of strategies designed to cope with the problems of message transmission and contextual change that arise in asynchronous communication. Just as the decision routine facilitates decision-making work, the use of epistolary forms and conventional structures in asynchronous interaction may allow participants to combine and minimize their managerial work.

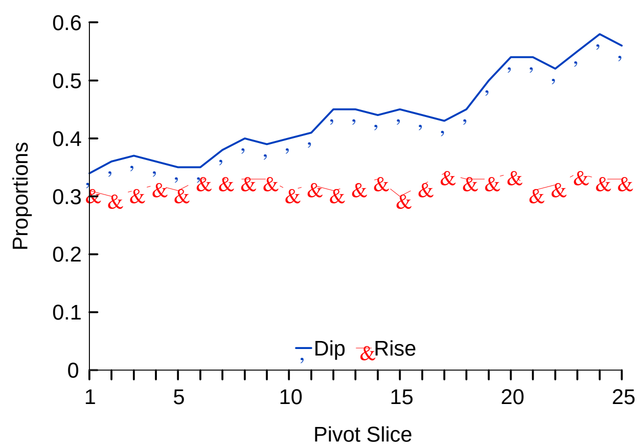
The lack of temporal co-presence in asynchronous interaction also unleashes a broad range of new behaviors not observed in the synchronous interactions. No longer constrained by the processing demands and memory limitations of synchronous interaction, participants have time to reflect, to edit and play with their messages, and to employ more elaborate encoding strategies. Since the long turns in the asynchronous interactions reflect these differences from the long turns in the synchronous interactions, a turn profile could usefully be designed in which the X-axis also reflects degrees of synchronicity in the interaction by spacing contributions corresponding to the delays between broadcasts. Then the profiles could be associated with synchronous short turn strategies, synchronous long turn strategies, and asynchronous long turn strategies, each with the organizing principles just described.

4. Turn profiles and discourse functions

Because every utterance in the interactions is annotated for discourse functions like suggesting and agreeing, we are able to explore links between the functions of utterances and superficial turn features such as turn size. In particular, we have focused on pivot turns, which we define as turns that are much shorter than the turns they precede or follow. Our data make it possible to estimate the probability that a pivot turn or the utterance before or after the pivot turn will exhibit a particular discourse function. For example, a short turn might be more likely to encode predictable, routine utterances such as agreements, which are typically encoded in minimal forms. Moreover, since agreements typically follow suggestions and precede orientations, we might see an increase in the likelihood that utterances before and after pivot turns will be coded as Suggests Action and Orients Suggestion respectively. Using data from the 4-line MTV condition, Figure 3 presents the cumulative proportions of pivot utterances coded as Agrees with Suggestion when the pivot utterance is defined by a dip (the pivot is shorter than the turn which precedes it) or by a rise (the pivot is shorter than the turn which follows it) of at least x words. In this figure, we systematically vary the difference in words between the pivot utterance and the utterance that precedes (for dips) or follows (for rises). This size, *the pivot slice*, is varied from 1 to 25 on the X-axis. Because these are cumulative functions, a pivot slice of 10 words, for example, includes all of the larger pivot slices, as well.

The proportion of pivot utterances coded as Agrees with Suggestion is higher than the average .11 proportion of utterances coded as Agrees with Suggestion in the 4-line corpus, and the proportion increases to almost .6 with increasing pivot slice, but only for pivots identified at dips. Thus utterances that are much shorter than the utterances that precede them are very likely to be utterances agreeing with the longer suggestions that precede them. The clear correlation between pivot slice size and proportion of Agrees with Suggestion illustrates how pivots, which are easily tracked by software, may be used to identify certain functions without having to do much on-line content analysis. We are currently working on incorporating pivot turn information in statistical models of the interactions such as those used for machine learning [4,13]. Other turn-based features may also prove to be predictive of discourse functions, too. For example, the location of an utterance as initial, medial, or final in a turn may be as significant in the synchronous turns as it is in the asynchronous messages where early utterances reflect opening strategies and final sequences encode closings. Finally, the extent to which turn features can be associated with discourse functions may depend on the type of turn strategy that participants select, and our data will make it possible to make this kind of refined comparison.

Figure 3: Cumulative proportions of utterances identified as pivots by dips and rises and coded as Agrees with Suggestion, varying the difference in



words between the pivot and the preceding (dips) or following (rises) utterance (=pivot slice)

5. Dominance and overlap

Because participants must share the channel in face-to-face interaction, their contributions can be viewed as competing for the limited resource of the floor. In this view, the relative size of participants' contributions reflects their success or dominance in competition. Though we are aware that refined measures of dominance should take into account the function of utterances along with the size [7], the amount of participants' contributions provides a content-free measure for comparative studies. Čech and Condon observe that participants tend to match each others' contribution size in the MTV corpora with average message sizes of partners achieving correlations of .52, .76 and .69 in the 4-, 10-, and 18-line conditions respectively [2]. Yet Figure 4 provides a turn profile in which the asymmetry of contributions is dramatically evident.

We have experimented with several ways of comparing participants' contribution sizes. By a simple count of the proportion of turns in which one partner had more to say than another, egalitarian dyads have an expected proportion close to .5. Another measure is composed of the ratio of the more verbose to the less verbose partner (defined over average number of words per turn: an expected value of 1.0 for egalitarian pairs) multiplied by the proportion of turns on which the more verbose partner had a longer message (thus an expected value of .5 for egalitarian pairs). The 4-line condition had a slightly higher levels of dominance than the 18-line condition on both of these measures (.681 versus .588 for proportion of turns, and 1.114 versus .870 for modified ratio), $F(1,38) = 6.95$, $MSe = .0125$, $p < .02$; $F(1,38) = 3.98$, $MSe = .1487$, $p < .054$. Therefore, it appears that the shorter turn strategies in the 4-line condition (see Table 2) provide a greater opportunity for one partner to dominate.

Interruptions and overlaps are sometimes used to measure dominance in interaction, but many factors affect these practices, and data from studies of computer-mediated interaction suggest that the communication environment is one of those factors. Van der Wege and Clark use measures of overlap to argue that turns are emergent properties of interaction because the average .03 proportion of words that overlap in a Unix Talk environment is significantly ($p < .005$) higher than the average .02 proportion that overlap in face-to-face and telephone conversations, and the number of overlapping words is higher in theTalk environment (5.71) compared to the face-to-face (5.16) and telephone (4.55) interactions ($p < .02$) [19]. Our data provide a more meaningful estimate of the extent to which overlap occurs because we compare proportions of utterances containing overlapped words rather than proportions of words. The utterance unit seems better for comparison here because utterance units reflect some of the organization and packaging of the communication, yet utterance units can be identified without the additional determination of their membership in turns.

Dividing the instances of overlap (counting one for each speaker) by the number of utterances in each interaction, we obtain an average proportion of .069 for the 16 face-to-face interactions in the first corpus and .063 for the 8 face-to-face interactions in the MTV corpus. In contrast, the interactions in Babineaux's corpus, obtained under conditions similar to the Talk environment, contain an average proportion of .22 overlapping utterances. Thus, our results resemble those of van der Wege and Clark, but with even higher incidence of overlap in the computer-mediated corpus. The larger estimate may result from software that makes identification of overlap more accurate by tracking the timing of each character typed by participants. If utterances are classified as overlapped when any character overlaps, more accurate identification of minimal overlaps will increase the proportions measured, especially with each counted twice.

We have suggested that measures of overlap can provide estimates of the degree to which ill-defined turn boundaries might influence results based on turn profiles. The magnitudes reported above indicate that the potential for this influence is generally small, though Babineaux's data suggest that overlap may present a larger problem in some corpora.

6. Environments that facilitate interaction

Comparisons of face-to-face and computer-mediated interaction have increased understanding of both types of interaction, but we are wary of the often tacit assumption of some researchers that face-to-face conversation reflects optimal engineering standards for verbal interaction. The research reported here shows that participants in interaction adapt their encoding strategies to the exigencies of the event, yet we found that participants did more than adapt to the novel synchronous computer-mediated environments. Some participants began to exploit the minimal asynchronicity of the system to produce long turns that are not ordinarily favored in synchronous environments. One participant sent a long turn, then immediately began to formulate another long turn that begins, "I did that so you could start reading while I keep typing." Therefore, researchers should consider the possibility that they may be underestimating human abilities if face-to-face interaction is assumed to reflect ideals or limits for all other language-based interaction.

In particular, it seems uncontroversial that rates of reading comprehension can greatly exceed rates of comprehension for spoken language, since the latter is limited by the rates at which people can produce spoken language and ultimately by the rates at which aural signals can be pro cessed. In contrast, not only are reading rates high compared to speaking rates, but they can be significantly improved with training. Moreover, reading makes it possible to apply processing strategies not possible in oral communication such as skipping over unimportant material. Yet reading comprehension rates would undoubtedly decrease if readers simultaneously formulated and positioned a next turn while remembering previously processed material, as they must in face-to-face interaction. Therefore, a design strategy to maximize comprehension speeds in synchronous computer-mediated environments could unbuckle production from comprehension and aim to reduce demands on memory and turn management.

A limited degree of asynchronicity in synchronous systems seems to reduce the need for participants to simultaneously comprehend the partner while planning their own next turns. This favors longer turns and the efficient reading strategies they afford. With speech recognition software to speed up production, such a system might exceed the speed and effectiveness with which words are processed in face-to-face interaction--at least in some contexts. Moreover, it seems clear that the benefits of enhancing reading comprehension rates will not be realized as much, if at all, when interaction is organized by short turn strategies. Therefore, the enhancements may be facilitative only for contexts in which long turn strategies are appropriate. This reasoning provides motivation for the hybrid synchronous/asynchronous systems that are being developed. Specifically, a system with a small window and minimal editing capabilities will suffice for short, fast turns, while a larger window with more extensive editing capabilities can provide for longer turns that could be inserted into synchronous interactions, e-mail or files (with customizable menus for frequently used listserves, mailing lists, or file locations).

Our results also motivate the efforts of researchers to identify ways of facilitating turn-taking. Devices to facilitate turn management might also include color-coding messages to identify the participants who produced them, or for environments with many participants, colored icons could identify message sources. In the latter environments, participants often produce *directed turns*, which are turns that begin by identifying a participant to whom the turn is particularly directed. Like the interpersonal introductions in Herring's data, directed turns perform interpersonal work and anchor the utterance to the common ground established by previous utterances. Directed turns could be facilitated by allowing participants to click on the screen name (or colored icon) of other participants to automatically add that participant's name as the beginning of a directed turn.

Finally, these considerations also support researchers' efforts to provide an easily searchable representation of previous discourse that can be used to relieve demands on short term memory. A graphic representation like the one provided in the Fugue system might be enhanced with color-coded segments that indicate message sources, which would help participants locate messages from particular participants or identify locations by sequences of authors. Clearly, there are many possibilities for representing and searching previous messages that designers can explore in their efforts to liberate synchronous communication from the limitations of working memory.

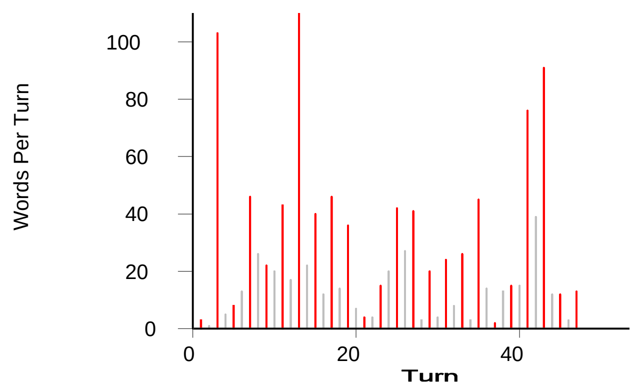


Figure 4: Profile illustrating dominance

7. Conclusions

The studies reported here underscore researchers' claims for the significance of turn organization in interaction. We can demonstrate that superficial properties of turns reflect fundamental organizing principles of the interaction which impact both the cognitive processing of the decisions in serial vs. parallel strategies and the encoding strategies that package this work. We also provide evidence that specific discourse functions may be linked to features of turns such as the pivot property and that interpersonal factors such as dominance can be reflected in turn profiles. A proposal to include temporal relationships in turn profiles rather than just sequential relationships not only would capture the degree of synchronicity in the interactions, but also could represent simultaneous utterances for those interactions that include them. Finally, turn profiles with accurate time measurements can be used to compare the processing rates of language under varying conditions.

The possibility of maximizing processing rates in computer-mediated communication raises issues of how efficiency and effectiveness of interaction can be measured. Clearly basic research is needed to obtain comparable measures in the many environments that can now be investigated along with face-to-face interactions. It will be especially difficult to take into account the interaction of communication strategies with other strategies that people use to organize their work. For example, we suspect that participants in the synchronous long turn interactions did not simply sit idle while waiting for their partner's messages. Comparisons of communication strategies should take into account other work accomplished simultaneously, and this multi-tasking potential is another human capability that computers have the potential to enhance. But again, basic research and comparisons like the ones we are making are called for.

A serious limitation of the results reported here is that they are based almost entirely on dyadic interactions. When we compare the 2- and 3-person e-mail interactions in Gray's study, participants in the latter formulate shorter turns from the start, as if following a tacit understanding that maintaining effective interactivity for 3 people would require shorter turns [10]. In fact, these results might be better explained by observing that in a 3-person interaction, the greater delays required to produce longer messages increase the chances that messages will be rendered redundant or irrelevant by messages sent by other participants in the meantime. The same considerations can account for Cherney's observation that turns in chat environments become shorter when the number of participants increases [3]. Consequently, we must be cautious in drawing conclusions from our results until additional data can be compared. Nevertheless, the results reported here suggest that people may be capable of a much broader range of communicative behaviors than researchers have anticipated.

8. References

[1] L. Babineaux, Simultaneous Transmission and Discourse Planning in Computer-Mediated Communication, M.S. Thesis, University of Louisiana at Lafayette, Forthcoming.

[2] C. Čech and S. Condon, "Message size constraints on discourse planning in synchronous computer-mediated communication", *Behavior Research Methods, Instruments, and Computers*, 30, 1998, pp. 255-263.

[3] L. Cherney, *Conversation and Community: Chat in a Virtual World,* Cambridge University Pr, Cambridge, 1999.

[4] W. Choi, J. Cho, and J. Seo, "Analysis system of speech acts and discourse structures using maximum entropy model", *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics,* Morgan Kaufman, San Francisco, 1999, pp. 230-237.

[5] S. Condon and C. Čech, Manual for Coding Decision-Making Interactions, Unpublihed 1992, Rev. 1995. Available via ftp://sls-ftp.lcs.mit.edu/pub/multiparty/coding_schemes/

[6] S. Condon and C. Čech, " Functional comparison of face-to-face and computer-mediated decision-making interactions", In S. Herring, Ed., *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives,* John Benjamin, Philadelphia, 1996, pp. 65-80.

[7] S. Condon and C. Čech, "Discourse management functions in three modalities", In S. Herring, Ed, *Computer-Mediated Conversation Analysis,* Forthcoming.

[8] S. Condon, C. Čech, and W. Edwards, " Measuring conformity to discourse routines in decision-making interactions", *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics,* Morgan Kaufman, San Francisco, 1999, pp. 238-245.

[9'] S. Herring, "Two variants of an electronic message schema". In S. Herring, Ed., *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives,* John Benjamin, Philadelphia, 1996, pp. 81-106.

[10] C. Gray, Does group size in electronic mail affect communication efficiency and flaming? M.S. Thesis, University of Louisiana at Lafayette, 1999.

[11] G. Jefferson, "A case of precision timing in ordinary conversation: Overlapped tag-positioned address terms in closing sequences", *Semiotica,* 9, 1973, pp. 47-96.

[12] M. McLaughlin, *Conversation: How talk is organized,* Sage Series in Interpersonal Communication, Vol. 3, Sage Publications Beverly Hills, CA, 1984.

[13] M. Nagata and T. Morimoto, "First steps toward statistical modeling of dialogue to predict the speech act type of the next utterance". *Speech Communication,* 15, 1994, pp. 193-203.

[14] H. Sacks, E. Schegloff, and G. Jefferson, "A simplest systematics for the organization of turn-taking for conversation". *Language* 50, 1974, pp. 696-735.

[15] E. Schegloff, "Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences", In D. Tannen, Ed., *Analyzing discourse: Text and talk,* Georgetown University Press, Washington, D.C., 1982, pp. 71-93.

[16] E. Schegloff, "Analyzing single episodes of interaction", *Social Psychology Quarterly,* 50(2), 1987, pp. 101-114.

[17] E. Schegloff, " Turn organization: One intersection of grammar and interaction", In E. Ochs, E. Schegloff, and S. Thompson, Eds., *Interaction and Grammar*, Cambridge University Press, Cambridge, 1996, pp. 52-133.

[18] T. Shankar, M. VanKleek, A. Vicente, and B. Smith, "Fugue: A Computer Mediated Conversational System that Supports Turn Negotiation", In IEEE *Proceedings of the Hawai'i International Conference on System Sciences,* Maui, Hawaii, 2000.

[19] M. Van der Wege and G. Clark, "Turn-taking systems for computer-mediated communication", In M. Shafto, and P. Langley, Eds., *Proceedings of the 19th Annual Conference of the Cognitive Science Society,* Lawrence Erlbaum, Mahwah, New Jersey, 1997, p. 1077.